



UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Detección y Modelado de Carriles de Vías Interurbanas mediante Análisis de Imágenes para un Sistema de Ayuda a la Conducción

Autor:

Juan Manuel Collado Hernáiz

Directores:

Arturo de la Escalera Hueso

Jose María Armingol Moreno

DEPARTAMENTO DE INGENIERÍA DE SISTEMAS Y AUTOMÁTICA

Leganés, Octubre de 2009

TESIS DOCTORAL

DETECCIÓN Y MODELADO DE CARRILES DE VÍAS INTERURBANAS MEDIANTE ANÁLISIS DE IMÁGENES PARA UN SISTEMA DE AYUDA A LA CONDUCCIÓN

Autor: Juan Manuel Collado Hernáiz

Directores: Jose María Armingol Moreno
Arturo de la Escalera Hueso

Firma del Tribunal Calificador:

Firma

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Calificación:

Leganés, de de

Los científicos nunca se cansan de decirnos que los frutos de su trabajo son “neutrales”; si enriquecen o destruyen a la humanidad depende de cómo son usados. ¿Y quién es el que decide cómo han de ser usados? No hay nada en la formación de los científicos e ingenieros que les permita tomar tales decisiones.

— Ernst F. Schumacher (*Lo pequeño es hermoso*, 1973)

A mi abuelo Juan.

1914 – 2007

Abstract

The goal of this work is the study and design of a Lane Departure Warning System. The system has been implemented on the research platform IVVI (“Intelligent Vehicle based on Visual Information”).

The system analyses road image sequences captured from an on-board video system. The aim is to anticipate dangerous manoeuvres and warn the driver, for instance, when the vehicle is about to cross a solid line, or before an unintended lane crossing.

This is achieved by means of the combined work of four algorithms. A first algorithm detects and tracks the road lines. A second algorithm classifies these lines into three different classes (solid lines, dashed lines, and merge lines). A third algorithm detects the lane change manoeuvres and warns the driver if the manoeuvre is dangerous or unintended. Finally, a self-calibration algorithm is able to calibrate the vision system without the need of artificial partners. The whole processing is performed in real time. The system has been tested under real conditions on the IVVI platform.

Resumen

El objetivo de la presente tesis es el estudio y diseño de un Sistema de Alerta por Salida de Carril, cuya implementación se ha realizado sobre la plataforma de investigación IVVI (*“Intelligent Vehicle based on Visual Information”*).

Este sistema analiza e interpreta secuencias de imágenes de carretera capturadas desde un sistema de percepción embarcado en el vehículo, con la finalidad de anticipar maniobras peligrosas y alertar al conductor, como por ejemplo, antes de cruzar una línea continua o de una salida inadvertida de carril.

Esto se consigue gracias al trabajo combinado de un algoritmo que detecta y sigue las líneas viales, otro que las clasifica en tres tipos distintos (continuas, discontinuas y de entrada/salida), y finalmente un tercero que detecta las maniobras de salida de carril y alerta al conductor ante maniobras peligrosas. Además, un cuarto algoritmo se encarga de autocalibrar el sistema de visión sin necesidad de patrones artificiales. Todo el procesado se efectúa en tiempo real. El sistema ha sido probado en condiciones reales sobre la plataforma IVVI.

Paso corto, vista larga, cachaza y mala intención.

— Refranero particular de mi abuelo Gregorio.

Agradecimientos

Esta tesis se debe a:

En primer lugar, mi familia más cercana. A mis padres y a mi hermana, por su educación, por su acompañamiento, por servir de guía y por enseñarme los valores de la solidaridad, la comprensión y el respeto, y por supuesto, el trabajo y la disciplina, imprescindibles en todo doctorando.

A mis abuelos. De todos ellos aprendí y aprendo. De mi abuelo Juan, el trabajo y la honestidad en todas sus acepciones. De mi abuela María, el optimismo y la risa. De mi abuela Rufina, la caridad y el servicio con mayúsculas, y de mi abuelo Gregorio, la humildad y el cariño.

A mis tíos Jesús, Enrique, Jose Carlos, Juan María, Fátima, Rosario, María Luisa, por su compañía día a día y verano a verano, y por sus consejos certeros.

A mis amigos, en especial aquellos que me han acompañado estos últimos años, Fernando, Isabel, Jose Alberto, Cecilia, Patricia, Luis, Samuel, Nacho, Álvaro, Jesús y Juan Pablo, y también, a mis viejos compañeros de piso, Jose Alberto, Cristina y Marta, a todos por estar ahí.

A mis compañeros de trabajo, Nando, Marco, Basam, Daniel, y a las nuevas incorporaciones María, Gustavo y Gema, por hacer del grupo un grupo.

A mis directores de tesis, Jose María y Arturo, y de grupo, Paco Pepe, por su apoyo constante, por su comprensión en los momentos fáciles y difíciles, por crear un ambiente de trabajo humano y agradable y por dejar en mi paso por la universidad un buen recuerdo.

Y a todas aquellas personas cuyas vidas se han cruzado con la mía en algún momento, dejando la impronta que me ha permitido ser lo que soy (entre otras muchas cosas, alguien capaz de realizar un doctorado).

Índice general

Índice de figuras	XII
-------------------	-----

Índice de tablas	XVIII
------------------	-------

1. Introducción	1
1.1. Motivación de la tesis	1
1.1.1. Estadísticas de accidentes	1
1.2. Sistemas Inteligentes de Transporte	7
1.2.1. Sistemas de percepción	9
1.3. Objetivos de la Tesis	9
1.4. Esquema del documento	10
2. Estado de la técnica	13
2.1. Introducción	13
2.2. Principales enfoques y dificultades	14
2.2.1. Esquema genérico de los sistemas de detección de carreteras	14
2.2.2. Principales enfoques	15
2.2.3. Principales dificultades	16
2.3. Sistemas de visión	17
2.4. Técnicas de pre-procesamiento de la imagen	20
2.4.1. Detección de regiones	21
2.4.2. Detección de bordes	23
2.4.3. Enfoques Mixtos	25
2.5. Técnicas de detección de carriles	26
2.5.1. Enfoques ascendentes	28
2.5.2. Enfoques descendentes	30
2.6. Proyectos clásicos	38
2.6.1. ALVINN (Carnegie Mellon University)	38
2.6.2. RALPH (Carnegie Mellon University)	40
2.6.3. ARGO (Universidad de Pavia)	43
2.6.4. UTA II (DaimlerChrysler)	45
2.7. Sistemas comerciales	47
2.7.1. AutoVue, de Iteris	47
2.7.2. LDW de TRW	48

2.7.3.	LDW de Mobileye	48
2.7.4.	Lane Departure Warning (LDW) / Lane Keeping System (LKS) de Continental AG	49
2.7.5.	HELLA - DAS	49
2.7.6.	AssistWare - SafeTRACK	50
2.7.7.	Toyota Lexus LS 460	50
2.7.8.	Irisbus - CIVIS	50
2.7.9.	Honda Accord	51
2.8.	Conclusiones	51
3.	Autocalibración del sistema de visión estéreo	55
3.1.	Introducción	55
3.1.1.	Parámetros de calibración	55
3.1.2.	Importancia de la calibración	57
3.1.3.	Trabajos Previos	58
3.1.4.	Breve Descripción	60
3.2.	Descripción del Algoritmo	61
3.2.1.	Modelo del Sistema de Visión Estéreo	62
3.2.2.	Obtención del patrón de calibración	65
3.2.3.	Función de Salud del Algoritmo Genético	70
3.3.	Resultados	73
3.3.1.	Con Imágenes Sintéticas	73
3.3.2.	Con Imágenes Reales	77
4.	Detección y seguimiento de carriles	81
4.1.	Introducción	81
4.2.	Perspectiva inversa	85
4.2.1.	Implementación de la perspectiva inversa	85
4.3.	Detección de marcas viales	86
4.3.1.	Ecualización de la imagen	88
4.3.2.	Cálculo del gradiente	88
4.3.3.	Búsqueda del patrón de marca vial	89
4.3.4.	Filtrado de ruido con operaciones morfológicas	90
4.4.	Modelo de carretera recta	92
4.4.1.	Detección de líneas viales con la Transformada de Hough	92
4.4.2.	Detección de líneas viales en Regiones de Interés (RDIs)	94
4.4.3.	Seguimiento de carriles con el filtro de Kalman	96
4.4.4.	Modificaciones en el mecanismo de votación de la Transformada de Hough	100
4.4.5.	Resultados	102
4.5.	Modelo de carretera con curvatura	106
4.5.1.	Seguimiento de carriles con un filtro de partículas	107
4.5.2.	Resultados	110

5. Clasificación de Líneas y Alerta por Salida de Carril	115
5.1. Clasificación de líneas viales longitudinales	115
5.1.1. Introducción	115
5.1.2. Análisis frecuencial de líneas viales	119
5.1.3. Algoritmo de Clasificación	126
5.1.4. Resultados	128
5.2. Sistema de Alerta por Salida de Carril	128
5.2.1. Reconocimiento automático del número de carriles . . .	128
5.2.2. Generación de alertas	130
5.2.3. Resultados	131
6. Conclusiones y Trabajos Futuros	135
6.1. Aportaciones de la tesis	136
6.2. Trabajos futuros	137
6.2.1. Mejoras en el Modelo de Carretera	137
6.2.2. Mejoras en el Sistema de Adquisición de Imágenes . . .	138
6.2.3. Integración de la Detección de Carreteras con el Resto de Módulos de Detección del IVVI	138
A. Plataforma de Experimentación	139
A.1. El Vehículo IVVI	139
A.2. Capacidades Sensoriales	142
A.2.1. Detección de Peatones	142
A.2.2. Detección de Señales de Tráfico	142
A.2.3. Detección de Vehículos	144
A.2.4. Reconocimiento de la cara del conductor	145
A.2.5. Detección de los Carriles de la Carretera	145
B. Rectificación de imágenes estéreo	147
B.1. Introducción	147
B.2. Modelo de cámara en coordenadas homogéneas	149
B.3. Matriz de proyección de perspectiva y geometría epipolar . . .	151
B.4. Calibración del sistema estéreo	154
B.5. Rectificación	156
C. Modelo parabólico de la carretera	161
C.1. Desarrollo en serie de Taylor de la circunferencia	161
C.2. Modelo de Carril parabólico	162
Bibliografía	165

Índice de figuras

1.1. Accidentes con víctimas en carretera y zona urbana (DGT, 2006a)	2
1.2. Tipos de accidentes en carretera y zona urbana (datos de 2006)	3
1.3. La luminosidad en los accidentes (datos de 2006)	4
1.4. Los factores atmosféricos en los accidentes (datos de 2006) . .	5
1.5. Accidentes en función de las características de la vía (datos de 2006)	6
1.6. Accidentes en función del estado del vehículo (datos de 2006) .	7
2.1. Esquema genérico de detección de carreteras	15
2.2. Sistema estéreo instalado en el IVVI	18
2.3. Sistema CeDaR, en (Dankers y Zelinsky, 2004)	18
2.4. Sistema de Visión Trinocular	19
2.5. MarVEye. Fuente: www.unibw.de	19
2.6. Visión omnidireccional, en (Zhu y otros, 1996).	20
2.7. Modelos de curvatura vertical.	32
2.8. ALVINN, en (Pomerleau, 1993).	40
2.9. Muestreo de la imagen, con inversión de perspectiva, en (Pomerleau y Jochem, 1996).	41
2.10. Estimación de la curvatura, en (Pomerleau y Jochem, 1996). .	42
2.11. Estimación de la distancia lateral, en (Pomerleau y Jochem, 1996).	42
2.12.	47
2.13. Sistema AutoVue de Iteris Inc.	48
2.14. Sistema SafeTrack, de AssistWare	50
2.15. Irisbus IVECO Civis	51
2.16. Algunos de los Sistemas ADAS del Honda Accord	51
3.1. Tipos de distorsión radial	56
3.2. Parámetros extrínsecos	56
3.3. Sistema Estéreo	57
3.4. Diagrama de flujo	61
3.5. Descripción del Algoritmo	62
3.6. Modelo del sistema de visión estéreo rectificado	63
3.7. Cambio de Sistema de Referencia	66

3.8. Detección de líneas viales para la generación del patrón de calibración	67
3.9. Detección de marcas viales basada en el gradiente	68
3.10. Efecto de la perspectiva sobre la detección de marcas viales . .	69
3.11. Parametrización de Hough: $\rho = x \cos \vartheta + y \sen \vartheta$	69
3.12. Patrón de calibración; (+) patrón desde la cámara derecha; (o) patrón desde la cámara izquierda	71
3.13. Modelo de carretera proyectado sobre las cámaras del sistema de visión estéreo	73
3.14. Resultados del AG; (+) puntos proyectados desde la cámara derecha; (o) puntos proyectados desde la cámara izquierda; (-) puntos proyectados según los parámetros reales	75
3.15. Resultados con ruido gaussiano añadido a las imágenes sintéticas	76
3.16. Detección de líneas y generación del patrón en el primer fotograma de la secuencia de imágenes reales.	78
3.17. Resultado del Algoritmo Genético en el primer fotograma de la secuencia de imágenes reales.	78
3.18. Parámetros extrínsecos a lo largo de una secuencia	79
4.1. Organigrama General	81
4.2. Dificultades en la detección de líneas viales (I).	82
4.3. Dificultades en la detección de líneas viales (II).	83
4.4. Dificultades en la detección de líneas viales (y III).	84
4.5. Etapa 2: Inversión de perspectiva	85
4.6. Transformación de perspectiva	86
4.7. Etapa 3: Detección de marcas viales	87
4.8. (a) imagen remuestreada, a vista de pájaro; (b) candidatos a marca vial; (c) líneas detectadas por la Transformada de Hough.	87
4.9. Ajuste del histograma de la imagen.	88
4.10. Gradientes	89
4.11. Perfil de un corte transversal en la imagen original y en el gradiente.	90
4.12. Etapas de la detección de marcas.	91
4.13. Etapa 4: Detección y seguimiento de carriles.	92
4.14. Modelo de línea vial con parametrización $\rho-\theta$ ($\rho = x \cos \theta + y \sen \theta$)	92
4.15. Funciones de votación para la transformada de Hough.	93
4.16. Detección de líneas viales con la transformada de Hough; (a) vista desde la cámara; (b) perspectiva inversa; (c) detalle del acumulador de Hough	94
4.17. Modelo de carretera utilizando cuatro regiones de interés. . . .	95
4.18. Detección de líneas viales con y sin Regiones de Interés	95

4.19. Detección impidiendo cruces entre líneas; (arriba) sin Regiones de Interés; (abajo) con Regiones de Interés y detección de la orientación mayoritaria de las líneas; (c) y (f) muestran los acumuladores de Hough rotados 90° respecto a las figuras anteriores; (f) presenta además la proyección sobre el eje angular a la derecha, que corresponde a la orientación mayoritaria de las líneas.	97
4.20. Modelo de carril recto	98
4.21. Modelo de carretera	98
4.22. RDIs delimitadas por Kalman	101
4.23. Proceso de detección de carriles, mostrado sobre el acumulador de Hough.	102
4.24. Problemas con el ajuste del modelo de carril en las curvas. . .	103
4.25. Cálculo de los pesos w_i	103
4.26. Pesos para $\Delta y = 0,1$ m/píxel, $\vartheta = 0,05$ rad, y $H = 1,18$ m. . .	103
4.27. Ejemplos de detección de carriles. En cada ejemplo, la imagen en perspectiva inversa aparece junto a la esquina superior derecha. En la imagen original la zona remuestreada aparece delimitada por una línea blanca.	104
4.28. Detección de carriles con Kalman a lo largo de una secuencia .	105
4.29. Modelo de carril con curvatura	106
4.30. Imágenes utilizadas en el seguimiento del carril con el filtro de partículas.	109
4.31. Sistema de fijación de las cámaras al parabrisas	111
4.32. Secuencia correspondiente a la carretera de enlace de Leganés con la salida 28 de la autovía M-40, en Madrid.	111
4.33. Algunos ejemplos de la secuencia ilustrada en la figura 4.32. Los puntos finos representan líneas continuas, los gruesos y de mayor espaciado representan líneas de entrada/salida, y los de tamaño y espaciado medios corresponden a líneas discontinuas (ver capítulo 5).	113
5.1. Líneas discontinuas.	117
5.2. Líneas de entrada/salida.	118
5.3. Líneas continuas.	118
5.4. Fase de la línea vial M-1.1.	120
5.5. Influencia de la fase en el espectro de frecuencias, para la marca M-1.1.	121
5.6. Influencia del ruido en el espectro de frecuencias (I).	122
5.7. Influencia del ruido en el espectro de frecuencias (y II).	123
5.8. Señales de partida para el análisis en frecuencias, y resultado de la detección y clasificación de líneas (en rojo las <i>continuas</i> , en verde las <i>discontinuas</i> y en azul las de <i>entrada/salida</i>). . .	124

5.9. Espectro de potencia del perfil de intensidades en niveles de gris; (<i>izquierda</i>) perfiles de las líneas extraídos de la figura 5.8b, de izquierda a derecha; (<i>derecha</i>) densidad espectral de potencia	125
5.10. Espectro de potencia del perfil de intensidades en binario; (<i>izquierda</i>) perfiles de las líneas extraídos de la figura 5.8c, de izquierda da derecha; (<i>derecha</i>) densidad espectral de potencia	125
5.11. Espectro de potencia para las líneas viales de la figura 5.8. Las frecuencias características, así como los umbrales en magnitud (las líneas horizontales) se distinguen por el color: verde para las <i>discontinuas</i> , y azul para las de <i>entrada/salida</i>	127
5.12. Ejemplo de clasificación con tres carriles	129
5.13. Ejemplo de clasificación con dos carriles.	129
5.14. Ejemplo de clasificación errónea provocada por la curvatura vertical.	129
5.15. Ejemplo de clasificación errónea provocada por defectos de pintura.	129
5.16. Algunos ejemplos de carretera.	130
5.17. Predicción del cruce de la línea.	131
5.18. Ejemplos de alertas por salida de carril	132
5.19. Ejemplos de falsas alertas. La mayoría de las falsas alertas se producen en salidas o incorporaciones a la autovía (casos (a), (b) y (c)), porque el carril tiende a seguir la línea continua de la derecha. Las alertas erróneas se producen por una mala clasificación de la línea vial (en el caso (d) la provoca un vehículo que oculta un tramo de la línea).	133
A.1. IvVI	141
A.2. Interfaces del IvVI.	142
A.3. Detección diurna de peatones.	143
A.4. Detección nocturna de peatones.	143
A.5. Sistema de Reconocimiento de Señales de Tráfico	144
A.6. Detección de vehículos	145
A.7. Reconocimiento del estado de somnolencia	146
B.1. Disposición “ideal” de un sistema estéreo.	147
B.2. Necesidad de procesar las imágenes de un sistema estéreo; (a) y (b) Imágenes izquierda y derecha originales; (c) Mapa de disparidades entre (a) y (b).	149
B.3. Rectificación	149
B.4. Modelo de cámara.	150
B.5. Interpretación de las filas de la matriz de proyección de perspectiva.	152
B.6. La línea epipolar.	153
B.7. Posición de las líneas epipolares.	153

B.8. Calibración del sistema estéreo; (a) sistema; (b) y (c) patrón de calibración; (d) posiciones de la cámara respecto al patrón; (e) imágenes tomadas con la cámara izquierda; (f) errores. . .	154
B.9. Distorsión; (a) radial; (b) tangencial	155
B.10. Patrón de calibración y sistemas de coordenadas	156
B.11. Resultados; (a)(b) Imágenes izquierda y derecha sin rectificar; (c)(d) Imágenes izquierda y derecha rectificadas; (e) Mapa de disparidades entre (a) y (b); (f) Mapa de disparidades entre (c) y (d).	158
C.1. Modelo parabólico de una línea vial	163
C.2. Modelo parabólico de carril	164

Índice de tablas

3.1. Parámetros intrínsecos	74
3.2. Parámetros del AG	74
3.3. Resultados del AG	74
3.4. Más ejemplos de resultados del AG	74
3.5. Resultados con ruido gaussiano añadido a las imágenes sintéticas.	77
5.1. Tipos de marcas viales longitudinales según la Norma 8.2-IC .	116
5.2. Frecuencias y umbrales en la clasificación de líneas viales . . .	127
B.1. Valores de la calibración	156

Capítulo 1

Introducción

1.1. Motivación de la tesis

Los automóviles constituyen el medio de locomoción más utilizado en la actualidad. Según un estudio de la Dirección General de Tráfico ([DGT, 2006b](#)), en España el turismo es el vehículo más presente en los hogares, con una media de 1,46 turismos por hogar, existiendo al menos uno en el 89 % de los hogares, y siendo utilizado más de tres días a la semana por el 78 % de los conductores. Se trata por lo tanto de un medio de locomoción de uso masivo, cuyos beneficios son indudables, y el precio a pagar es la congestión del tráfico, el elevado número de accidentes, la contaminación, etc. Durante el año 2006 se produjeron en España 99.797 accidentes de tráfico con víctimas (35 accidentes por cada 10.000 vehículos del parque), en los cuales perdieron la vida 4.104 personas ([DGT, 2006c](#)). Por otro lado el número de vehículos no para de crecer. En España el parque se ha incrementado en un 19 % en el periodo 2000–2005. Este aumento hace que las infraestructuras se queden pequeñas, provocándose congestiones y falta de fluidez en ellas.

El objetivo de los Sistemas Inteligentes de Transporte (SIT) es incrementar la seguridad, eficiencia y confort del transporte mejorando la funcionalidad de los coches y las carreteras, usando las tecnologías de la información. Dentro de este conjunto de objetivos, esta tesis se centra en la seguridad, por lo que continuaremos este capítulo de introducción con un breve análisis de los accidentes de tráfico.

1.1.1. Estadísticas de accidentes

A continuación se presentan algunas estadísticas de accidentes de tráfico, extraídas del Anuario de Accidentes de la D.G.T. de 2006 ([DGT, 2006a](#)). Todas las estadísticas del informe son referidas a los accidentes de tráfico en los que hubo alguna víctima de las cuales se consideran mortales aquellas que murieron en las 24h siguientes al accidente y como consecuencia de éste.

Accidentes según el tipo de vía

La figura 1.1 representa el número de accidentes y de víctimas tanto en carretera como en zona urbana. Las cifras nos muestran que, a pesar de haber un mayor número de accidentes en zona urbana, la gravedad de los accidentes en carretera es mucho mayor, tanto en número de víctimas mortales como no mortales.

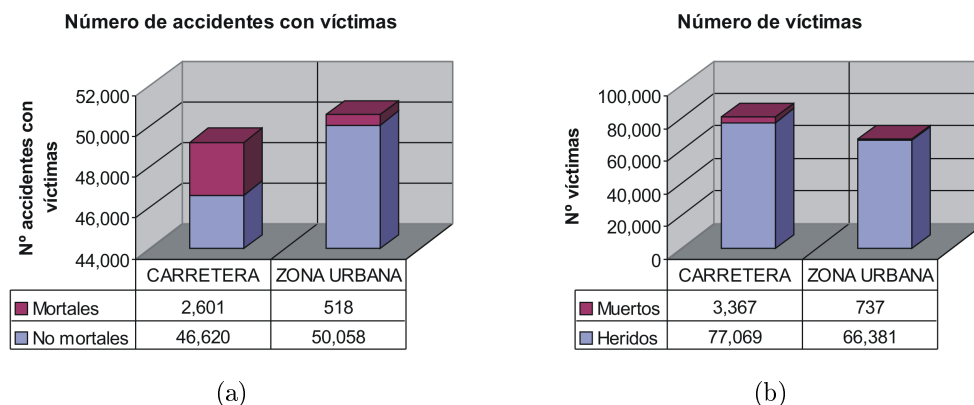


Figura 1.1: Accidentes con víctimas en carretera y zona urbana (DGT, 2006a)

Principales tipos de accidentes

La figura 1.2 muestra los principales tipos de accidentes mortales y no mortales en carretera y zona urbana. En las gráficas se aprecia que en zona urbana la mayoría de los accidentes se producen por colisión entre vehículos, mientras que la mayoría de las muertes se producen por atropello de peatones.

En carretera, en cambio, son la colisión entre vehículos en marcha y la salida de calzada las que acaparan el 87% de los accidentes y el 80% de las muertes. El sistema desarrollado en esta tesis pretende reducir el número de víctimas por salida de calzada en carreteras, que en 2006 provocó 1.191 muertos, 5.821 heridos graves, y 21.129 heridos leves. Si se comparan estas cantidades con el número de muertes por accidente de avión, que en España ha sido de 2.000 en los últimos 30 años¹, podemos hacernos una idea de la gravedad de estas cifras.

Factores del entorno en los accidentes

Luminosidad: Las figuras 1.3a y 1.3b representan los accidentes con víctimas en función de la luminosidad de la vía, y la luminosidad en función de la tipología de accidente, respectivamente. Según estos gráficos la mayoría de

¹<http://www.retraso.com/web/tusderechos/accidentes.htm>

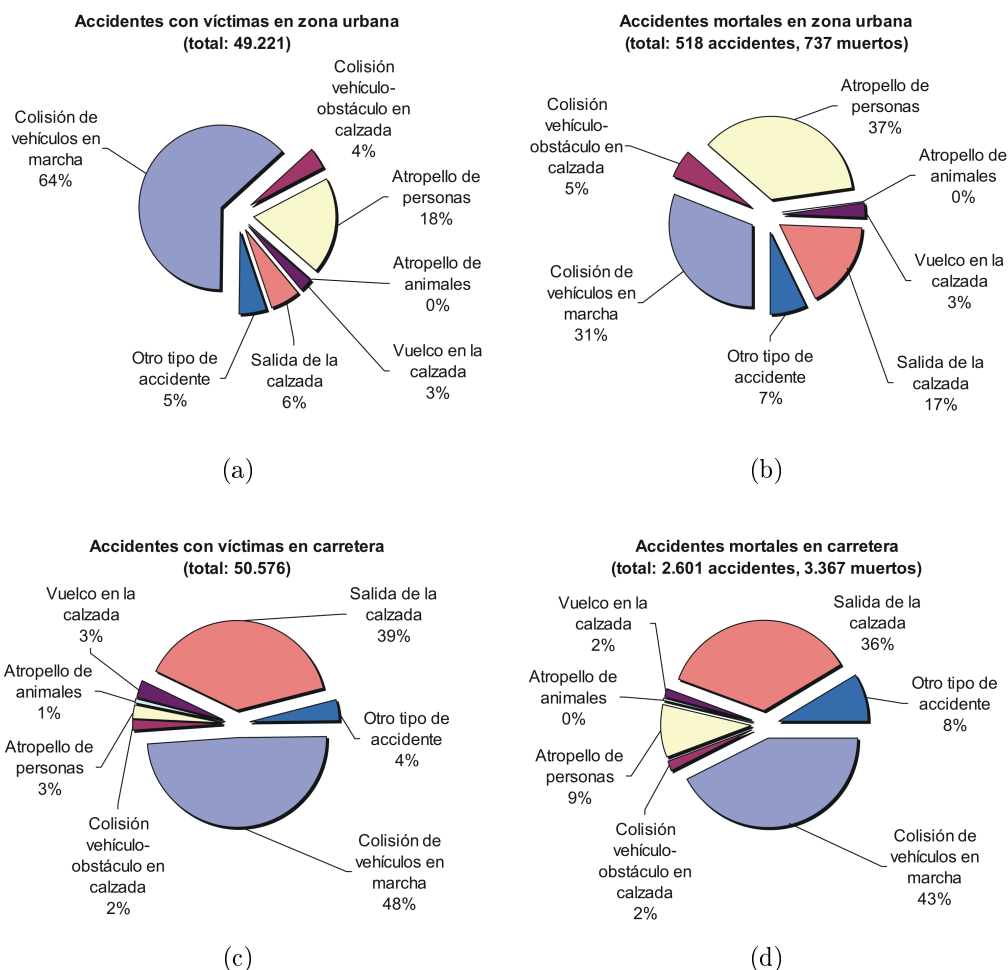


Figura 1.2: Tipos de accidentes en carretera y zona urbana (datos de 2006)

los accidentes en 2006 tuvieron lugar en pleno día (el 66,1 %), seguidos, aunque de lejos, por los producidos de noche en vías no iluminadas (el 15,8 %). Por tipo de accidente, los más frecuentes fueron por colisión entre vehículos (23.927), y por salida de la calzada (19.346), ambos en pleno día.

Factores atmosféricos: Atendiendo a los factores atmosféricos, resulta significativo el hecho de que en 2006 la gran mayoría de los accidentes (el 85,8 %) se produjeron con buen tiempo (figura 1.4).

Características de la vía: Los datos de accidentes por tipo de vía también arrojan cifras interesantes (figura 1.5). En 2006, la mayoría de los accidentes se produjeron en tramos rectos (47,2 %) o curvas suaves (16,3 %).

Estado del vehículo: Por último, sorprenden también los datos de accidentes según el estado del vehículo (figura 1.6), puesto que en la inmensa

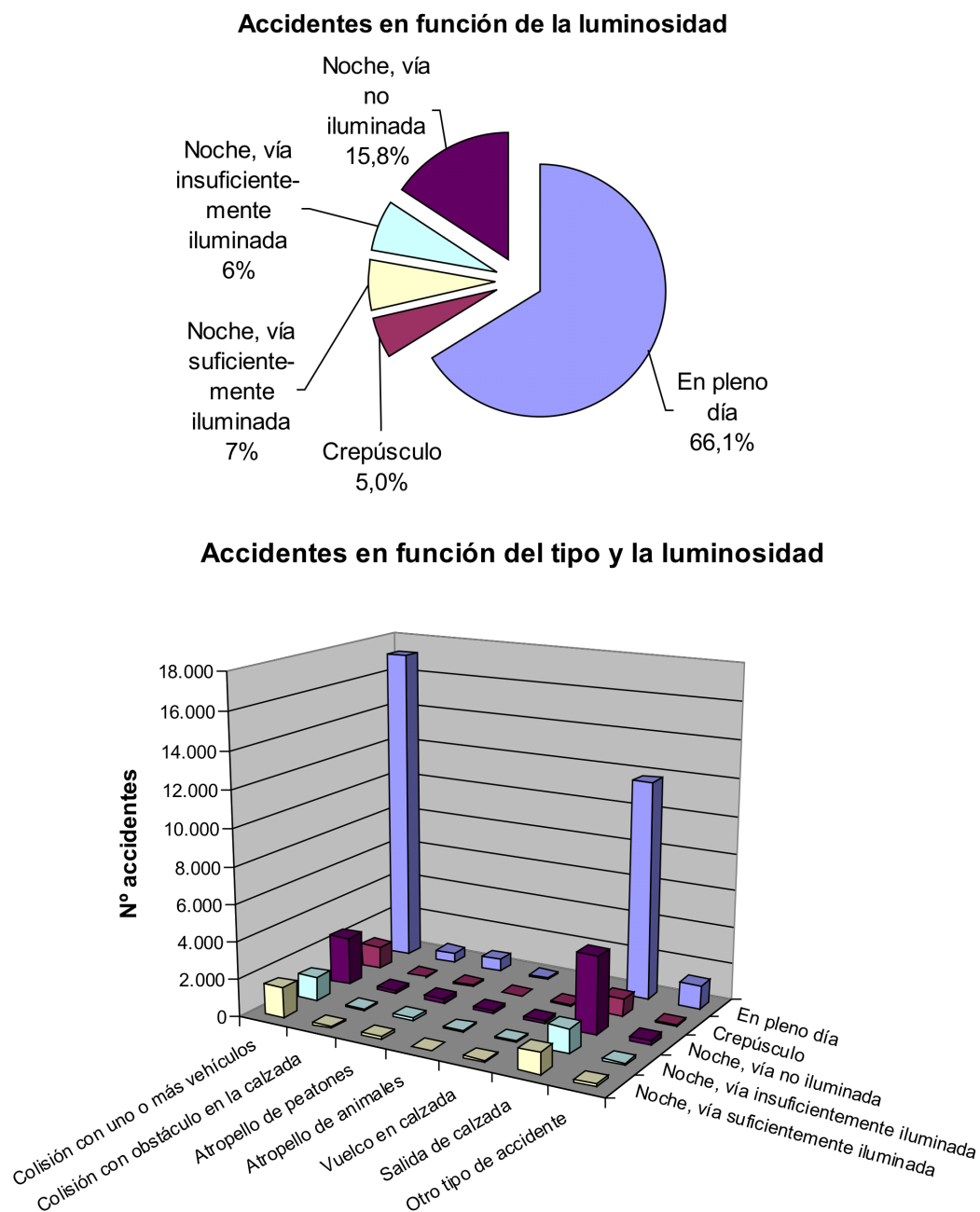


Figura 1.3: La luminosidad en los accidentes (datos de 2006)

mayoría de los accidentes intervinieron vehículos sin ningún defecto aparente (97,5 %).

En este punto cabe recordar que las estadísticas no establecen por sí solas relaciones causa-efecto, y es necesario interpretarlas como lo que son, pues de lo contrario se podría caer en el absurdo de pensar, por ejemplo, que el buen tiempo y la ausencia de defectos en los vehículos son algunas de las

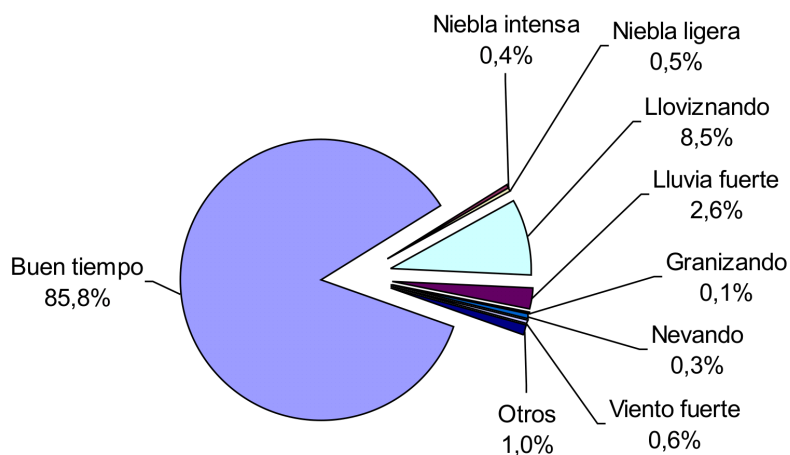
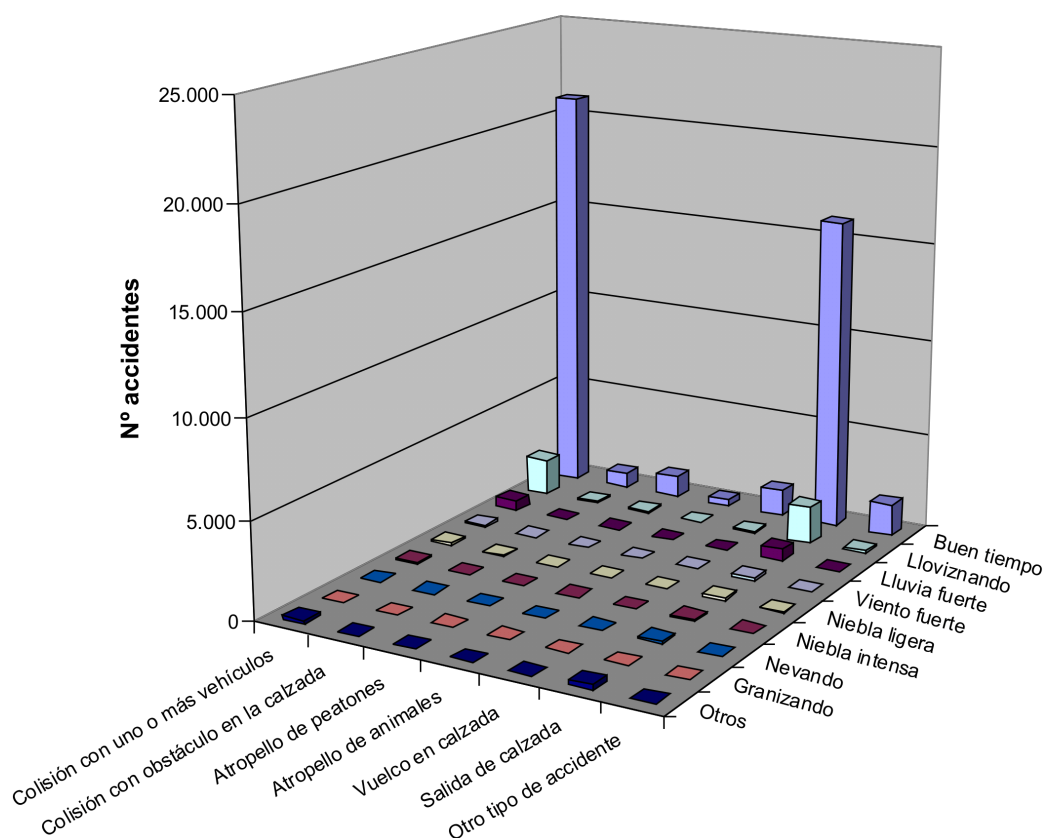
Accidentes en función de los factores atmosféricos**Accidentes en función del tipo y los factores atmosféricos**

Figura 1.4: Los factores atmosféricos en los accidentes (datos de 2006)

principales causas de accidentes. La mayoría de desplazamientos de vehículos se producen en condiciones favorables (de día, con buen tiempo, etc.), por lo que no resulta sencillo establecer conclusiones sobre la influencia del entorno

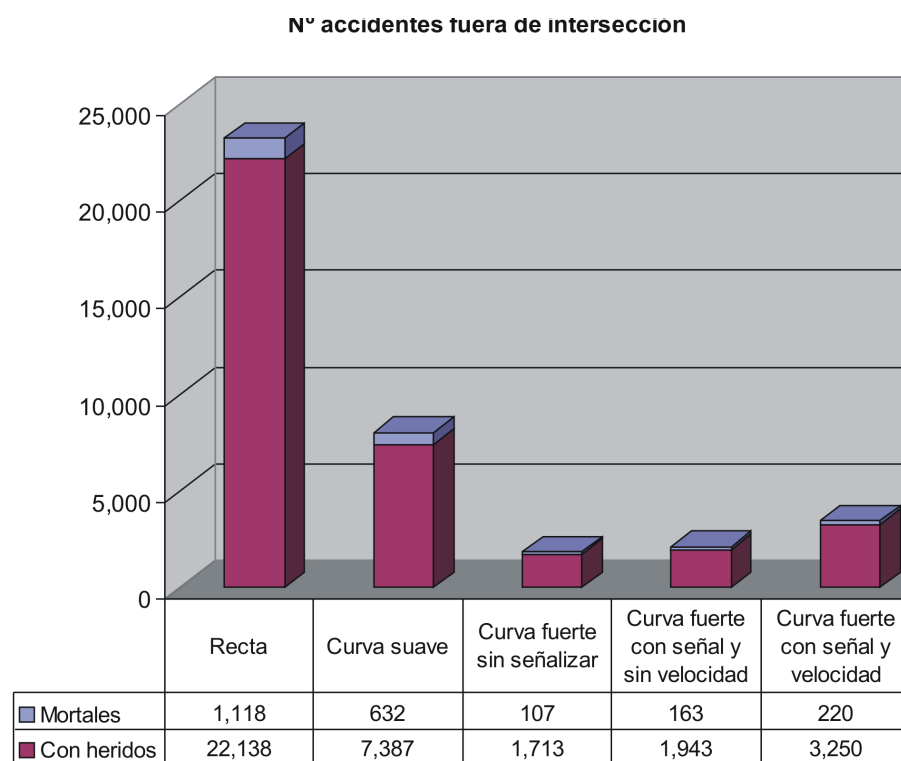
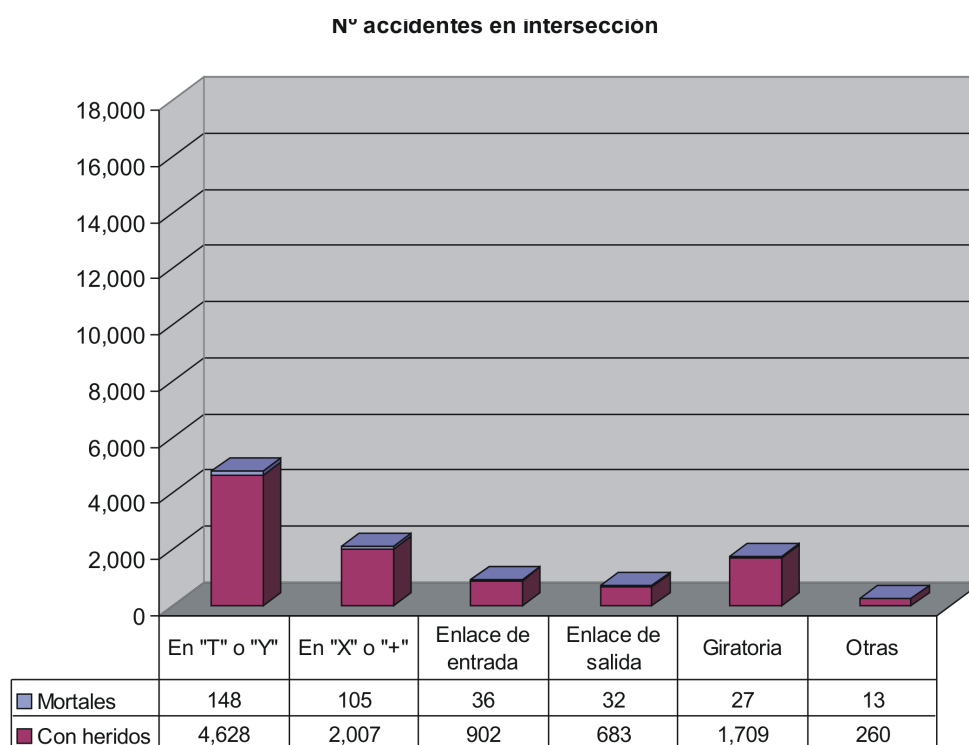


Figura 1.5: Accidentes en función de las características de la vía (datos de 2006)

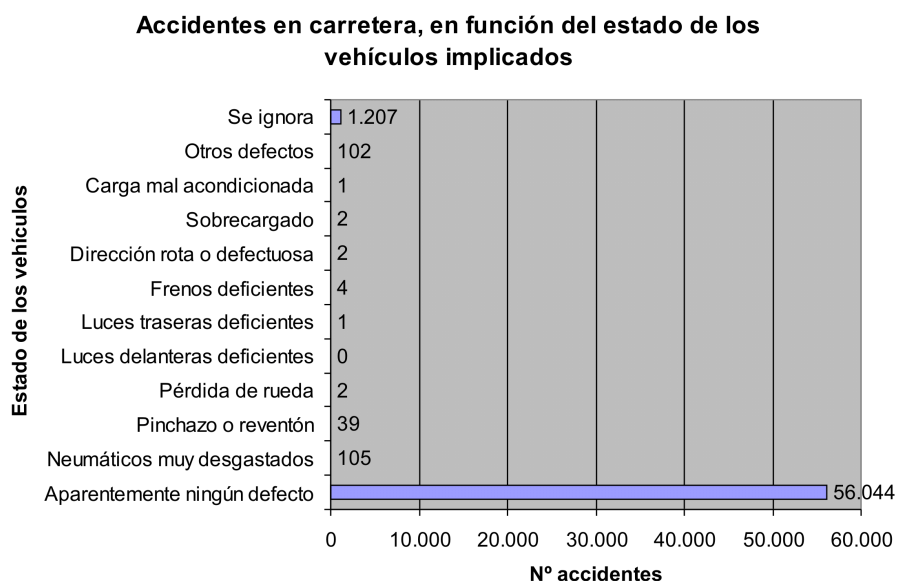


Figura 1.6: Accidentes en función del estado del vehículo (datos de 2006)

en la siniestralidad en la carretera.

Sin embargo, estos datos sí parecen apuntar que no hay una relación directa ni evidente entre las causas del accidente y los factores ambientales, por lo que se advierte una notable influencia del factor humano. Para caracterizar este factor humano, sería necesario un estudio más profundo que la mera consulta de estas estadísticas, pero el hecho de que la mayoría de los siniestros se produzcan en condiciones favorables (de día, con buen tiempo, en rectas o curvas suaves, y con vehículos en buen estado), parece apuntar que es el exceso de confianza una de las principales causas de los accidentes de tráfico en carretera.

1.2. Sistemas Inteligentes de Transporte

En este apartado se presentan algunas de las técnicas que, dentro de los Sistemas Inteligentes de Transporte (SIT), pretenden reducir el número de accidentes y paliar sus consecuencias.

Existen dos campos en los que se puede trabajar dentro de los SIT:

- mejoras introducidas en las infraestructuras,
- dotar a los vehículos de nuevas capacidades.

Ambas soluciones tienen sus ventajas e inconvenientes. La primera supone una buena opción si se trata de rutas pequeñas y para desplazamientos fijos de vehículos públicos, pero introducirlas en la red viaria total de un

país presentaría un coste prohibitivo. Por ello, en la actualidad parece más razonable poner el énfasis en los vehículos. Más aún si, como se ha visto en el apartado 1.1.1, las estadísticas ofrecidas por la DGT parecen sugerir que es el fallo humano la principal causa de los accidentes.

Motivados por lo anterior, han surgido una serie de campos en los que se está trabajando para lograr una conducción automática:

- Seguimiento del borde de la carretera.
- Mantenimiento de la distancia de seguridad.
- Regulación de la velocidad dependiendo del estado del tráfico y del tipo de la carretera.
- Adelantamientos de otros vehículos.
- Trazado automático de la ruta más corta.
- Movimiento y aparcamiento dentro de las ciudades.

Sin embargo, aunque los beneficios de una conducción completamente automática son numerosos, presenta varias dificultades:

1. **Técnicas:** Los algoritmos que realicen estas tareas deben trabajar en tiempo real con un nivel de fiabilidad próximo al 100 %.
2. **Económicas:** Su incorporación no debe suponer un incremento significativo del coste actual de los vehículos.
3. **Psicológicas:** Los ocupantes de los vehículos deben acostumbrarse a no controlar la marcha de éstos y ser conducidos por un ordenador.
4. **Legales:** En caso de accidente, la delimitación del alcance de las responsabilidades del conductor y del fabricante del vehículo es compleja.

Por estos motivos, y persiguiendo el objetivo de incrementar la seguridad en el transporte, se ha optado por englobar esta tesis dentro del campo de los Sistemas Avanzados de Ayuda a la Conducción, si bien, siendo conscientes de que la tecnología desarrollada puede ser aplicada, al menos parcialmente, a la conducción automática en el futuro. El que al final estas tecnologías sean un paso intermedio hacia la conducción automática o no, dependerá sobre todo de cómo se resuelvan los problemas técnicos, legales y psicológicos.

Dentro de estas ayudas a la conducción, los sistemas más importantes para un vehículo son:

1. **Sistema de aviso en caso de adormecimiento** (en inglés, “*Drowsy Driving Warning System*”). Monitorizan el estado del conductor y le avisan en caso de que se produzca una pérdida de atención sobre la conducción.

2. **Control de velocidad variable** (“*Adaptive Cruise Control*”). Se adapta la velocidad a la del vehículo que hay enfrente manteniendo la distancia de seguridad accionando el acelerador y el freno.
3. **Sistema anti-colisión** (“*Anti Collision Assistant*”). Parecido al anterior pero en el que los obstáculos son objetos estáticos, como vehículos parados, objetos en la vía, etc. Se avisa al usuario de la presencia de obstáculos, o coches detenidos o a velocidades muy bajas.
4. **Parar y marchar** (“*Stop & Go*”). Mantienen el control del vehículo a bajas velocidades, por ejemplo en atascos o en colas para entrar en las autopistas, en semáforos o en peajes.
5. **Sistema de ángulos muertos** (“*Overtaking Warning*”). Los sensores cubren el ángulo muerto del vehículo avisando de la presencia de otros vehículos que estén realizando un adelantamiento.
6. **Sistema de alerta por salida de carril** (“*Lane Departure Warning*”). El sistema detecta de forma automática la posición respecto a la línea lateral avisando al conductor si la va a sobrepasar de forma inadvertida.

1.2.1. Sistemas de percepción

El sistema de percepción de los SIT puede basarse en sensores pasivos como las cámaras, o activos como radares y láseres. Aunque estos últimos funcionan mejor que las cámaras en condiciones ambientales adversas, la visión por computador presenta una serie de ventajas frente a ellos:

- La mayoría de los accidentes se producen de día y con buen tiempo (buena visibilidad).
- Los radares y láseres sólo pueden obtener una representación muy simple del entorno, basada en distancias. Por otro lado, al ser sistemas activos, es necesario solucionar el problema de las interferencias entre los sensores de varios vehículos cuando su uso se generalizase. En el caso de los radares, su baja resolución hace que los obstáculos próximos entre sí se detecten como uno solo.

Si bien, los sensores activos pueden captar información del entorno en situaciones en las que los sensores pasivos fallan, como por ejemplo en condiciones de muy baja visibilidad, de modo que un uso combinado de ambos sistemas cubriría un mayor rango de situaciones.

1.3. Objetivos de la Tesis

Se plantea como objetivos de esta tesis, el desarrollo de un sistema de ayuda la conducción basado el análisis de imágenes con los siguientes criterios:

- Que sea capaz de detectar la realización de maniobras arriesgadas o incorrectas, en concreto las relacionadas con las salidas o cambios de carril que se realizan de forma no intencionada, o que aún siendo intencionadas incumplen la normativa o implican un riesgo de accidente.
- Que sea de fácil instalación en el vehículo, en concreto, debe facilitar en lo posible la calibración del sistema.
- Que suponga un coste asequible dentro del sector del automóvil.

Estas capacidades requieren de una serie de habilidades:

- a) Reconstrucción del perfil de la carretera.
- b) Determinación de la posición del vehículo dentro de la vía.
- c) Identificación de los distintos tipos de líneas viales (continuas, discontinuas de separación de carriles normales, discontinuas de separación de carriles de entrada/salida, etcétera).
- d) Identificación del tipo de carretera (autopista, autovía, secundaria, etcétera) y el número de carriles.
- e) Estimación de la fiabilidad de las observaciones, para poder detectar de forma automática en qué casos el sistema no es capaz de interpretar las observaciones.
- f) Alertar al conductor en situaciones de maniobras incorrectas o peligrosas.
- g) Autocalibración del sistema, sin necesidad de recurrir a patrones artificiales.

Las habilidades (c), (d) y (g) constituyen tres de las aportaciones novedosas de esta tesis. Estas novedades podrán apreciarse mejor tras la lectura de los capítulos 2 y 3.

1.4. Esquema del documento

Este documento comienza con este capítulo de introducción donde se han definido el marco y los objetivos de la tesis, como respuesta a una problemática social. En el capítulo 2 se presenta un estudio de otros trabajos realizados en esta línea de investigación, que servirá para enmarcarla dentro del campo tecnológico de los Sistemas de Ayuda a la Conducción en la actualidad.

Los siguientes capítulos detallan cada uno de los cuatro algoritmos diseñados en esta tesis. El primero de ellos, presentado en el capítulo 3, es el sistema de autocalibración, cuya finalidad es obtener de forma automática

los parámetros extrínsecos (posición y orientación) del sistema de visión, sin necesidad de un patrón de calibración artificial que complicaría en exceso la instalación en el vehículo.

El capítulo 4 detalla los dos sistemas desarrollados de detección y seguimiento de carriles, uno que implementa un modelo recto, y otro que implementa un modelo curvo. Este último empezó como una evolución del modelo recto, pero acabó siendo un nuevo desarrollo con técnicas de procesamiento y seguimiento diferentes a las del modelo anterior.

El capítulo 5 describe el sistema de clasificación de líneas viales, que, combinado con el de detección y seguimiento, ha permitido definir un sistema de alerta por salida de carril con capacidades que no se encuentran actualmente en prácticamente ningún sistema similar, como la distinción entre el cruce de una línea continua y una discontinua.

En el capítulo 6 se exponen las conclusiones y se proponen una serie de líneas de investigación que se derivan directamente del trabajo aquí presentado.

Por último, en el apéndice A se define la plataforma experimental (el vehículo IVVI) que ha permitido realizar pruebas a los algoritmos diseñados en condiciones reales, y se detallan las características del vehículo, así como el resto de sistemas que se están actualmente desarrollando en él. El apéndice B describe las ecuaciones y el proceso de rectificación de imágenes estéreo, que proporciona las imágenes de partida de todos los algoritmos aquí desarrollados.

Capítulo 2

Estado de la técnica

2.1. Introducción

Los primeros sistemas de detección de carreteras basados en visión por computador comienzan a aparecer en la década de los 80, y estaban fundamentalmente orientados a la conducción automática. En EEUU destaca la Carnegie Mellon University (CMU), con sus sistemas SCARF y YARF (Thorpe y otros, 1991) para el vehículo NavLab financiados por la agencia de investigación para la defensa DARPA (*“Defense Advanced Research Projects Agency”*), y en Europa la Universidad de Bundeswer-Munchen con su sistema VITA (Dickmanns y Mysliwetz, 1992), para el vehículo VaMoRs, dentro del proyecto europeo PROMETHEUS.

Sin embargo, en la última década están surgiendo cada vez más iniciativas dedicadas al diseño de sistemas de ayuda a la conducción, motivados por las razones comentadas en el apartado 1.1 del capítulo 1. Aunque ambos enfoques comparten la misma tecnología, existen ligeras diferencias entre ellos:

- En la conducción automática se extrae únicamente la información necesaria para deducir el giro del volante, la velocidad, etc. El objetivo del sistema de percepción es proporcionar una descripción del entorno que permita la navegación autónoma del vehículo. Puede ser suficiente con estimar la posición y orientación instantánea del vehículo dentro del carril, incluso con un modelo muy sencillo de carretera, pero la fiabilidad de estos datos debe ser muy alta. En la actualidad, la mayor parte de la investigación en este campo se orienta hacia fines militares.
- En la asistencia al conductor, los requisitos de percepción son diferentes. Ya no se trata de tomar el control del vehículo, sino de percibir todos los elementos relevantes del entorno para evaluar el potencial riesgo de cada situación. Por lo tanto, es necesario obtener más información del entorno para evaluar la situación y predecir maniobras. Si se está a tiempo de evitar el accidente, el sistema deberá emitir las alertas necesarias para que el conductor ejecute una acción correctora, y si el

accidente es inevitable, se deberá preparar el vehículo para minimizar las consecuencias. En general, la investigación en este campo persigue fines comerciales y/o sociales.

Por ello, aunque ya se han conducido vehículos de forma semi-automática durante largas distancias, donde cabe dar especial mención a los clásicos sistemas RALPH (Pomerleau y Jochem, 1996), VaMP (Maurer y otros, 1996) y ARGO (Broggi y otros, 1999), en los últimos años ha surgido un creciente interés por los Sistemas de Asistencia a la Conducción, tanto en el ámbito científico como en el social. Cada vez son más las noticias aparecidas en los medios de comunicación de masas, y los sistemas mostrados en ferias del automóvil.

2.2. Principales enfoques y dificultades

Dentro de los Sistemas de Ayuda a la Conducción, la detección de carreteras con un sistema de visión embarcado en un vehículo en movimiento, y en ambientes exteriores, se enfrenta, entre otras, a dos grandes dificultades:

- Por una parte, los cambios bruscos de iluminación, así como la presencia de suciedad, sombras, brillos, reflejos, grietas, parches de asfalto y otros obstáculos en la carretera dificultan en gran medida el tratamiento e interpretación de la imagen, así como el modelado del entorno.
- Por otra parte, la necesidad de dar respuestas en tiempo real requiere que los algoritmos utilizados sean de gran eficiencia. Más aún si, como se menciona en el capítulo anterior, se desea utilizar equipos convencionales de visión y procesamiento para evitar que el sistema global adquiera un alto precio. Esta condición es importante si se quiere que el sistema tenga una utilidad social.

2.2.1. Esquema genérico de los sistemas de detección de carreteras

En general, los algoritmos de detección de carreteras siguen un esquema similar al de la figura 2.1, que se compone de las siguientes etapas:

1. Se comienza por la *captura de la imagen* mediante alguno de los sistemas de visión que se comentarán en el apartado 2.3.
2. A continuación, en la etapa de *pre-procesado*, se resaltan las características que se consideran más representativas de la carretera o de las líneas viales, como por ejemplo los bordes, el color, la textura, etc. Generalmente el resultado de esta etapa es una imagen en la que aparecen segmentados los objetos que son candidatos a pertenecer a marcas

viales o a la calzada. Algunos algoritmos incluyen una etapa de validación, en la que mediante ciertos criterios se rechazan o confirman las observaciones.

3. En la etapa de *post-procesado*, bien se agrupan los objetos ya detectados para formar estructuras más complejas, o bien se ajusta un modelo predefinido, que puede ser paramétrico o no, a las observaciones de la etapa anterior. Algunos algoritmos incluyen también técnicas de seguimiento como el filtro de Kalman o el de partículas.
4. Finalmente, se utiliza la información anterior para una *aplicación* concreta, como puede ser el seguimiento automático de carril, las alertas por salida de carril, u otros sistemas relacionados con la asistencia a la conducción.

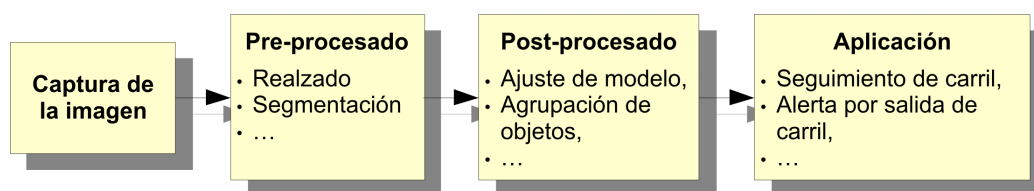


Figura 2.1: Esquema genérico de detección de carreteras

2.2.2. Principales enfoques

Dentro de la detección de carreteras para vehículos inteligentes, podemos distinguir tres grandes bloques de métodos: los basados en técnicas de aprendizaje, como por ejemplo las redes neuronales, los basados en la extracción de características de la imagen (enfoques ascendentes), y por último, los basados en el ajuste de un modelo (enfoques descendentes).

1. **Basados en aprendizaje:** Son métodos que pueden llegar a ser muy eficientes en tareas muy concretas. Sus principales inconvenientes son, por una parte, la necesidad de un conjunto representativo de patrones de entrenamiento, nada fácil de conseguir –por ejemplo, en (Pomerleau, 1993) se tiene que recurrir a un pequeño artificio para ampliar artificialmente el conjunto de patrones (ver apartado 2.6.1)–, y por otra, la todavía limitada capacidad de generalización en un entorno tan variable como el de carretera. Por ello, habitualmente estos métodos se aplican, bien en la etapa de *pre-procesamiento* exclusivamente, para segmentar los objetos de interés, bien para *conducción automática*. En este último caso se suele aplicar una red neuronal para procesar y obtener la información necesaria para el cálculo del giro del volante. No hay necesidad de interpretar la imagen ni de reconstruir el perfil de la

carretera. Dentro de este bloque destaca el proyecto ALVINN, que por su especial relevancia en este campo de la investigación, se describirá en el apartado 2.6.1.

2. **Basados en características:** Se denominan también *enfoques ascendentes*, porque parten de las observaciones obtenidas en el pre-procesado (por ejemplo, pequeños segmentos), y las van agrupando entre sí, siguiendo unas reglas predefinidas para formar estructuras de más alto nivel (por ejemplo, un carril). Tienen como ventaja la capacidad de detectar todos los objetos que aparecen en la imagen (podrían, por ejemplo, detectar todas las líneas viales visibles en una imagen, y no sólo las que delimitan el carril propio). En cambio, estas técnicas se ven muy afectadas por oclusiones parciales o totales, y pueden ser confundidos por patrones que presenten una apariencia similar a la de los objetos buscados (por ejemplo, bajo ciertas condiciones de iluminación los guardarraíles se asemejan mucho a una línea vial).
3. **Basados en modelos:** También denominados *enfoques descendentes*, aludiendo a que parten de un modelo predefinido que se ajusta a las observaciones, generalmente mediante la minimización de una función de coste. Este modelo incluye el conocimiento *a-priori* que se tiene del mundo. Su inconveniente es que sólo pueden detectar aquellos objetos que se contemplan en el modelo (por ejemplo, no pueden detectar la presencia de dos carriles si el modelo sólo considera uno), pero tienen como ventaja la capacidad de localizar objetos parcialmente ocluidos (por ejemplo, pueden localizar el carril aunque sólo se vea uno de sus bordes).

2.2.3. Principales dificultades

Ya se comentó en la introducción de este capítulo que las dos grandes dificultades a las que se enfrenta la detección de carreteras son el procesamiento en tiempo real y la variabilidad del entorno.

El requisito de tiempo real exige técnicas o suposiciones que facilitan la detección de objetos y aceleran el proceso. Las técnicas más utilizadas para reducir el tiempo de procesamiento son:

- el análisis de regiones de interés (comúnmente denominadas *ROIs*, siglas en inglés de “*Regions Of Interest*”),
- las aproximaciones geométricas del entorno, como por ejemplo:
 - la carretera es plana (o localmente plana),
 - los carriles tienen un ancho constante (o localmente constante),
 - no hay curvatura vertical, o sigue un modelo preciso,

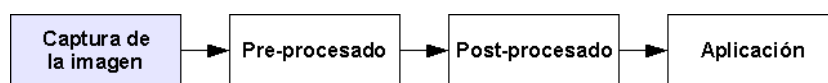
- etc.
- el remuestreo de la imagen para reducir la resolución,

Por otro lado la gran cantidad de ruido típico de los entornos exteriores dificulta enormemente la creación de sistemas robustos. Una forma de eliminar su efecto es procesar únicamente ciertas regiones de interés, evitando así el efecto del ruido que queda fuera de esas zonas. Estas regiones suelen ser pequeñas ventanas cuyo tamaño y posición se deducen de una etapa de predicción basada en el modelo de carretera, o bien son simplemente regiones que resultan de engordar los objetos detectados en la imagen anterior (Richter y Wetzel, 1995), al suponer que su posición no varía bruscamente de imagen a imagen. También se puede utilizar visión estéreo para excluir del procesado las regiones con objetos elevados sobre el plano de la carretera, y que por tanto no pertenecen a ella (Risack y otros, 1998; Taylor y otros, 1996; Bertozzi y Broggi, 1996).

Es también frecuente hacer un filtrado temporal para facilitar el seguimiento de los objetos. Para esto existen básicamente dos opciones: el *filtro de Kalman* (Bar-Shalom y Fortmann, 1988), y el *filtro de partículas* (Doucet y otros, 2001). En particular, dentro del campo de la visión por computador se emplea con frecuencia el algoritmo CONDENSATION (Isard y Blake, 1998). El filtro de partículas es una técnica más lenta que el filtro de Kalman, pero permite trabajar con distribuciones estadísticas diferentes de la normal y recuperarse más fácilmente cuando se pierde la pista del objeto a seguir.

2.3. Sistemas de visión

En este apartado se comentan los sistemas de visión más utilizados, que se corresponden con la primera etapa de procesamiento de los Sistemas de Asistencia a la Conducción:



Los sistemas más empleados en este campo de la investigación son:

- **Monocular:** Se suelen utilizar cámaras monocromas cuando se trabaja con enfoques basados en bordes, o a color cuando se trabaja con enfoques basados en regiones o mixtos (apartado 2.2.2). Generalmente se coloca cerca del espejo retrovisor interior del vehículo. La ventaja es que se tiene menos información que procesar que los sistemas multicámara. Los sistemas monoculares no permiten realizar una reconstrucción completa del entorno en 3D, por lo que obligan a recurrir a hipótesis geométricas como las suposiciones de carretera plana o carriles de ancho constante.

- **Binocular o Estéreo:** Permite obtener información de profundidad. Cuando se utiliza aplicado a la detección de carreteras, se suele emplear para detectar y excluir del análisis de la imagen aquellas regiones que contengan obstáculos o para facilitar el ajuste de modelos de carretera más complejos, como aquellos que tienen en cuenta la curvatura vertical (Nedevschi y otros, 2004). En entornos urbanos también se ha utilizado para localizar los límites de la calzada mediante la detección del bordillo de la acera (Turchetto y Manduchi, 2003). A su vez, los sistemas estéreo pueden ser:
 - *Montura fija:* La posición relativa de ambas cámaras es fija. Una vez calibrado el sistema, mediante el proceso de rectificación se obtiene una pareja de imágenes que proceden de cámaras virtualmente idénticas. Es el sistema más común (figura 2.2).



Figura 2.2: Sistema estéreo instalado en el IVVI

- *Montura móvil:* Las cámaras pueden variar su orientación de forma independiente. Permite orientar las cámaras a voluntad para seguir objetos o realizar estéreo con diferentes configuraciones, pero requiere un sistema control complejo y preciso. Un ejemplo es el sistema CeDaR (figura 2.3) (Dankers y Zelinsky, 2004).

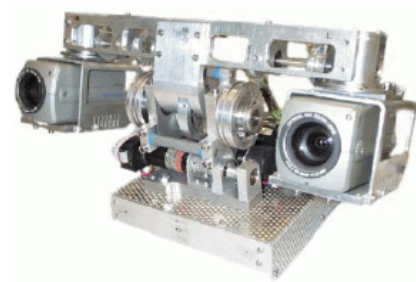


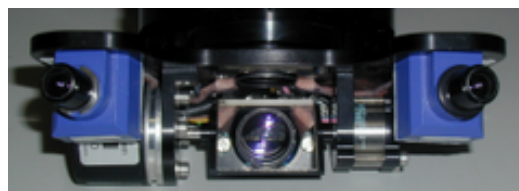
Figura 2.3: Sistema CeDaR, en (Dankers y Zelinsky, 2004)

- **Trinocular:** Sistema formado por tres cámaras dispuestas en ángulo recto (figura 2.4), de forma que posibilita la aplicación de técnicas estéreo tanto en horizontal como en vertical. Un ejemplo de aplicación se describe en (Williamson y otros, 1999).



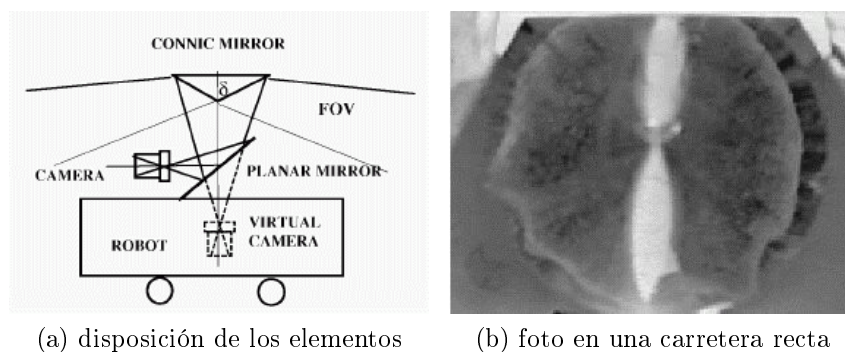
Figura 2.4: Sistema de Visión Trinocular

- **“Multi-focal”**: Para cubrir la pérdida de resolución de los objetos lejanos al vehículo, en relación a los cercanos, se han desarrollado sistemas que combinan cámaras con diferentes distancias focales, para adaptarse tanto al entorno próximo al vehículo (con un gran angular) como al lejano (con un teleobjetivo). Un ejemplo de estos sistemas es el *MarVEye* (acrónimo en inglés de “*Multi-focal Active/Reactive Vehicle Eye*”) (Dickmanns y Mysliwetz, 1992) que consiste en un conjunto de cámaras de distintas características cuya configuración, posición y orientación está gobernada por un sistema de control. El *MarVEye* permite enfocar simultáneamente varias zonas de la imagen –a distintas distancias–, y así modificar dinámicamente el campo de visión de forma que aquellos objetos que se quiere analizar o seguir aparezcan lo mejor posible en la imagen. Los movimientos de las cámaras son bruscos, imitando al ojo humano, pero generan vibraciones que impiden el procesamiento durante dos o tres fotogramas. Requiere de un complejo sistema de control, denominado “*EMS-Vision*” (siglas en inglés de “*Expectation-based Multifocal Saccadic Vision*”) (Gregor y otros, 2002). Otros autores, como Guiducci (1999), también sugieren la conveniencia de este tipo de sistemas de visión.

Figura 2.5: MarVEye. Fuente: www.unibw.de

- **“Omni-direccional”**: Son sistemas que permiten obtener fotogramas que cubran un campo de visión de 360° (Zhu y otros, 1996; Corke y otros, 2003; Trivedi y otros, 2007). Consisten en un espejo esférico o

cónico colocado en posición vertical, y una cámara dispuesta como en la figura 2.6a. El espejo proyecta una imagen que abarca un campo de visión lateral completo. El sistema es interesante, a la hora de detectar intersecciones, pero da una imagen difícil de tratar, y de baja resolución especialmente en el entorno cercano al vehículo.



(a) disposición de los elementos (b) foto en una carretera recta

Figura 2.6: Visión omnidireccional, en (Zhu y otros, 1996).

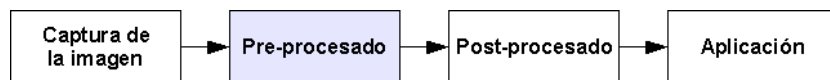
Los sistemas de visión para asistencia a la conducción deben ser adecuados para diversas aplicaciones (detección de carreteras, peatones, vehículos, señales de tráfico, etc.). De ahí que hayan surgido sistemas como el *CeDaR* o el *MarVEye*, de elevada versatilidad, pero que añaden una complejidad al control y a la calibración, además de un coste al sistema global. Los sistemas de visión omni-direccional ofrecen una baja resolución en el entorno cercano, y dificultan el procesamiento de las marcas viales, ya que la conversión entre los sistemas de coordenadas del mundo y de la imagen es compleja.

Para el caso del vehículo IVVI, en el cual se encuadra la presente tesis, se ha optado por un sistema estéreo de montura fija, que está siendo utilizado para las aplicaciones de detección de carreteras, vehículos y peatones.

Los algoritmos de detección de carreteras descritos en esta tesis sólo precisan un sistema monocular, excepto el algoritmo de autocalibración, descrito en el capítulo 3, que aplica técnicas estéreo para obtener los parámetros extrínsecos del sistema de visión –imprescindible para la detección de carreteras– de forma automática y sin necesidad de un patrón de calibración artificial.

2.4. Técnicas de pre-procesamiento de la imagen

En la etapa de *pre-procesamiento*, que comparten tanto los métodos basados en características como los basados en modelos, se extraen de la imagen las características más representativas de la carretera.



Dentro de esta etapa se pueden diferenciar tres enfoques distintos:

- *Basados en regiones*: Segmentan los píxeles pertenecientes a la carretera o a las marcas viales mediante la detección de regiones en la imagen, empleándose generalmente el color o la textura como características discriminantes.
- *Basados en bordes o en gradiente*: Utilizan la imagen de gradiente de intensidades u otros mecanismos de detección de bordes para extraer los bordes de la calzada o de las líneas viales.
- *Enfoques mixtos*: Combinan los dos anteriores.

Aunque todos estos enfoques pueden aplicarse tanto a la detección de la carretera como de las líneas viales, en el caso de detección de carretera, la mayoría de los trabajos tienden a un enfoque basado en regiones, mientras que en el caso de detección de líneas viales se suele tender hacia enfoques basados en bordes, como veremos a continuación.

2.4.1. Detección de regiones

La detección de regiones busca la clasificación de cada píxel en dos clases: perteneciente o no perteneciente al objeto que se quiere segmentar, ya sea la calzada de la carretera o las marcas viales.

Las técnicas basadas en regiones han sido aplicadas mayoritariamente a la conducción automática en entornos rurales ([Turk y otros, 1988](#); [Crisman y Thorpe, 1991, 1993](#); [Lin y Chen, 1991](#); [Sotelo y otros, 2004](#); [Franke y otros, 2007](#)), y son de gran utilidad en situaciones en que las marcas viales no están presentes. La detección de marcas viales requiere de mayor precisión que la de carretera, dado que el tamaño y forma de las líneas viales en la imagen las hace fácilmente confundibles con otros objetos (por ejemplo, grietas o surcos en el asfalto, guardarrailes, sombras alargadas, etc.). Sin embargo, dado que el Sistema de Asistencia propuesto en esta tesis necesita conocer la presencia y posición de los carriles, en este apartado nos centraremos en las técnicas aplicadas a la detección de marcas viales.

Como en todo problema de clasificación, es necesario determinar qué información o características son suficientemente representativas como para poder ser utilizadas como discriminantes. La característica más evidente es el *color*:

- Existen sistemas, como YARF ([Kluge, 1993](#)) o el propuesto en ([Sun y otros, 2006](#)), que segmentan las marcas viales por separación en el espacio de color HSI, y otros como ([Chiu y Lin, 2005](#)), que segmenta

umbralizando en un espacio definido por las diferencias entre cada par de componentes RGB. En general estos métodos son muy rápidos, pero no son capaces de abarcar todas las condiciones de iluminación que se producen en entornos exteriores –amanecer, día, atardecer, noche–.

Mientras los métodos anteriores utilizan *información global*, es decir, umbralizan por igual en toda la imagen, en otros trabajos se utiliza *información local*, de forma que cada píxel se clasifica en función de un entorno próximo a él:

- En (Gonzalez y Ozguner, 2000) se procesa la imagen por ventanas de anchura igual a la de la imagen y 5 píxeles de altura. Supone que la banda inferior corresponde sólo a la carretera, por lo que su histograma será unimodal, y los límites superior e inferior de la distribución corresponden a dos umbrales que delimitarán tres clases (obstáculo, carretera, y marca vial). La ventaja es que se adapta muy bien a condiciones de iluminación variables en el tiempo, pero no en el espacio, como ocurre de noche, o ante la presencia de sombras, y además requiere que la parte inferior de la imagen esté libre de obstáculos o manchas.
- En (Jeong y Nedevschi, 2003) se discrimina en función de la media del nivel de gris en un entorno cercano al píxel considerado, y de la media del nivel de gris en la imagen. Utiliza un clasificador basado en árboles de decisión cuya ventaja es la rapidez, pero no llega a obtener una segmentación precisa.
- En (Cheng y otros, 2006) se realiza primero una detección de carretera, en la que se basa para detectar después las líneas viales. Utiliza dos espacios de color diferentes, y un complejo sistema para determinar los umbrales, bajo la suposición de que los colores de las marcas viales siguen distribuciones gaussianas.
- Kim (2006) estudia varios clasificadores aplicados a los tres canales RGB de subventanas de 9x3 píxeles (en total, 81 elementos). Compara entre sí tres tipos de redes de neuronas, dos clasificadores bayesianos “naïve”, máquinas de soporte vectorial (SVMs) y el algoritmo propuesto en (Jeng y otros, 2003). En su estudio las máquinas de vectores soporte obtienen buenos resultados pero a un coste computacional inviable para aplicaciones de tiempo real, por lo que se decanta por las redes de neuronas.

En general, el uso exclusivo del color no suele conseguir una buena segmentación de las marcas viales en todos los casos. Además del color, la *textura* es la otra característica que se emplea con mayor frecuencia:

- En (Campbell y otros, 1997) se comparan los resultados de un clasificador basado en redes neuronales de base radial, utilizando primero

sólo color, y después, color más textura, concluyendo que el color no es suficiente para una correcta clasificación. A continuación calcula, para cada zona segmentada, 28 características que incluyen información de color, textura, forma, contexto y geometría, para clasificarlas con un perceptrón multicapa en 11 clases distintas que corresponden a objetos que típicamente se encuentran en entornos de carretera. Alcanza una precisión del 91 %, pero con un coste computacional considerable.

- Jeong y Nedeveschi (2005) utilizan el clasificador basado en árboles de decisión como etapa preliminar y a continuación aplican K-medias partiendo de la detección anterior, obteniendo mejores resultados que en (Jeong y Nedeveschi, 2003). Para acelerar los cálculos se realiza un procesado por subventanas.

Además del *color* y la *textura*, algunos autores han investigado el uso de un mayor número de discriminantes:

- En (Shanahan y otros, 1999) se han estudiado más de sesenta características, que incluyen información de color, posición, orientación, tamaño, forma y textura, para segmentar regiones mediante K-medias. Después utiliza un filtro basado en redes neuronales para seleccionar las diez características más representativas. Entre ellas aparecen las coordenadas horizontal y vertical del centroide de la región, es decir, el algoritmo reconoce que la carretera suele aparecer en la zona inferior de la imagen, y por lo tanto, la posición del píxel es relevante a la hora de clasificarlo. A continuación construye con programación genética un clasificador basado en lógica borrosa, lo que dota al algoritmo capacidades de auto-aprendizaje. Alcanza precisiones del 97 % y el autor menciona que es más rápido que otros algoritmos de clasificación, como los basados en Bayes, árboles de decisión, o redes neuronales.

El inconveniente que presentan estos enfoques es el elevado tiempo de procesado que requieren las técnicas de clasificación, en comparación con los métodos basados en bordes, de modo que su aplicación en sistemas de asistencia a la conducción tiende a realizarse en carreteras no marcadas o en entornos urbanos, donde las técnicas de detección de bordes son menos eficaces.

2.4.2. Detección de bordes

Las técnicas basadas en la detección de bordes necesitan de la existencia de unos bordes de la carretera bien marcados, ya sean los límites de la calzada o las marcas viales. Mientras que las técnicas de detección de regiones suelen aplicarse sobre el plano de la imagen, las técnicas de detección de bordes han venido siendo aplicadas también sobre la imagen obtenida tras invertir la relación de perspectiva, denominada imagen en *perspectiva inversa* (ver

apartado 4.2), y que muestra un plano cenital de la carretera. Aparte de esto, la variabilidad de técnicas basadas en bordes es mucho menor que las basadas en regiones, ya que prácticamente todas ellas se basan en el gradiente de intensidad de la imagen.

Cuando se quiere identificar no sólo las marcas viales sino también los límites de la calzada en carreteras no marcadas u otros objetos, se suele utilizar un detector de bordes convencional, sin tener en cuenta las características especiales que posean las líneas viales. El filtro de Canny (Canny, 1986) ofrece una buena relación señal-ruido y es utilizado con frecuencia. En (Foresti y Regazzoni, 1999) se aplica a la detección de bordes de la carretera, de vehículos o de peatones, ya que a continuación se pretende reconocer todos estos tipos de objetos en la misma imagen, y así obtener una interpretación del entorno. En (Wang y otros, 2000) se utiliza para detectar las líneas, si las hay, o los bordes de la calzada, cuando las líneas no están presentes.

Aunque es poco habitual, a la hora de detectar marcas viales sobre el plano de la imagen, es interesante tener en cuenta que la apariencia de las marcas, o de los bordes cambia con la distancia al vehículo. En (Luetzeler y Baten, 2000) se realiza un filtrado con una máscara que puede variar en altura, anchura y contenido, dependiendo de las características 3D esperadas para la carretera.

Cuando se pretende localizar exclusivamente las marcas viales, éstas se suelen modelar como una banda clara sobre fondo oscuro. En (Broggi, 1995) se realiza una inversión de perspectiva para obtener una imagen cenital de la carretera. Trabajar en este plano tiene la ventaja de que las marcas viales son fundamentalmente verticales y de anchura prácticamente constante. A continuación realiza un filtrado no lineal, por filas, para buscar transiciones claro-oscuro-claro. Para tratar de reducir la sensibilidad a las sombras, aplica una transformación morfológica que refuerza la intensidad de aquellos píxeles que cumplen condiciones de conectividad, y después binariza con un umbral adaptativo.

En (Ieng y otros, 2003) también se buscan transiciones oscuro-claro-oscuro, pero en el campo del gradiente, es decir, parejas de ascenso-descenso de gradiente que representan el borde izquierdo y derecho de una marca vial. A diferencia de (Broggi, 1995) lo hace en el plano de la imagen, lo que exige tener en cuenta que el ancho de marca varía con la coordenada vertical.

En (Kreucher y Lakshmanan, 1999) se presenta un enfoque completamente distinto. Realiza la detección de bordes en el dominio de la frecuencia. Analiza la imagen por pequeñas ventanas, aplicando la *Transformada Coseno Discreta* o *DCT* (*Discrete Cosine Transform*) en cada una de ellas. Estudiando los coeficientes obtenidos es capaz de determinar qué ventanas tienen una alta energía de bordes. A continuación se buscan bordes diagonales, ya que trabaja en el plano de la imagen, donde en un entorno próximo al vehículo las líneas de la carretera normalmente cruzan la imagen en diagonal hacia el punto de fuga. La ventaja del algoritmo es la rapidez. La desventaja

es que si la carretera es curva, en el entorno lejano los bordes pueden aparecer como horizontales o verticales.

Jung y Kelber (2005) aplican un paso adicional denominado “*Edge distribution function*” (EDF) para obtener bordes alineados en la dirección mayoritaria. Estos autores aplican el detector de Sobel, y a continuación construyen un histograma de orientaciones de bordes (EDF), donde suponen que las marcas viales corresponderán a máximos locales. El inconveniente de este método es que falla en presencia de sombras o bordes con direcciones similares a los de las líneas viales.

Otro trabajo en el que también se juega con las orientaciones de bordes es el de McCall y Trivedi (2006). En este trabajo se utilizan los denominados “*steerable filters*”, basados en segundas derivadas de gaussianas, y que pueden ser ajustados para detectar bordes en una dirección concreta. Estos autores lo aplican a dos casos diferentes, por una parte a la detección de líneas viales, que presentan una dirección de borde mayoritaria, y por otra a la detección de los reflectores que se colocan en ocasiones sobre las líneas viales, y que son visibles de noche al ser iluminados por los faros de los vehículos. Éstos últimos muestran unos bordes con una respuesta alta en todas las direcciones, ya que aparecen en la imagen como un círculo brillante. Según los autores la ventaja de estos filtros es su versatilidad.

Las ventajas de estas técnicas frente a las basadas en regiones son la rapidez, puesto que el cálculo del gradiente no es muy costoso, y la capacidad de detectar las líneas viales siempre que presenten un cierto contraste con la calzada. Un mínimo contraste puede ser suficiente para que las marcas viales puedan ser reconocidas aunque haya sombras, nubes, o sea de noche. Por lo tanto, estas técnicas fallan cuando las marcas tienen poco contraste con el asfalto, como sucede en condiciones de lluvia intensa, en las que las marcas son casi imperceptibles, o cuando la luz incide con un ángulo tal que todo el asfalto aparece brillante. También fallan cuando el contraste no es el esperado, como por ejemplo a la salida de los túneles o en ciertos atardeceres o amaneceres, cuando se produce el fenómeno de “inversión” del color, y el asfalto aparece más brillante que las marcas viales.

Otro inconveniente son las falsas detecciones que se producen en presencia de ruido estructurado¹, como fuertes sombras con bordes rectos, o estructuras como los guardarraíles que presentan bandas blancas muy similares a las líneas viales.

2.4.3. Enfoques Mixtos

Existen otros enfoques que combinan la detección de regiones con la de áreas. La idea es utilizar toda la información posible para elevar la robustez del algoritmo y reducir los falsos positivos:

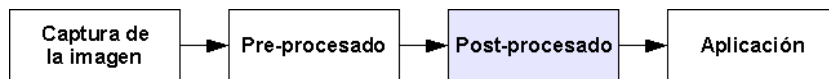
¹Nos referimos al ruido provocado por la presencia de otros objetos en la imagen, que interfieren con aquello que se pretende detectar

- En carreteras rurales, el traqueteo del vehículo da lugar a imágenes borrosas. Por eso en (Luetzeler y Baten, 2000) se utiliza la detección de bordes en entorno cercano, y la de área en entorno lejano.
- En (Apostoloff y Zelinsky, 2003) se utilizan una serie de *pistas*² que se fusionan mediante el filtro de partículas. Cada *pista* devuelve para cada píxel la probabilidad de que pertenezca a una marca vial. Como pistas utiliza una imagen de correlación con un modelo de marca vial (banda clara sobre fondo oscuro), una imagen de bordes obtenida con el detector de Canny, y el color que representa la zona de carretera.

La combinación de varias técnicas mejora la robustez, pero tiene el inconveniente de elevar mucho la carga computacional. Es necesario utilizar técnicas de optimización para su aplicación en tiempo real, así como evaluar si la mejora del rendimiento compensa el aumento de esta carga.

2.5. Técnicas de detección de carriles

La definición y ajuste de un modelo de carretera ha sido uno de los temas más trabajados en la investigación. Nos encontramos ahora en la etapa que en esta tesis se ha denominado *post-procesado*:



Entre los modelos clásicos se encuentran los tres comentados por Morgenthaler y otros (1990), que permiten realizar una reconstrucción tridimensional de la carretera con una sola cámara:

- **Carretera plana:** Este modelo asume que la carretera es plana, y que ese plano coincide con el que sustenta el vehículo. La ventaja de este método es la rapidez con la que se puede hacer la reconstrucción tridimensional de la carretera, ya que la relación entre los sistemas de coordenadas asociados a la carretera y a la imagen es lineal y viene definida por una matriz. Por otro lado, este modelo no tienen en cuenta la curvatura vertical de la carretera y se ve muy afectado por desviaciones en los parámetros extrínsecos de calibración de la cámara, en especial, por la variación en la inclinación de la cámara respecto de la carretera, por ejemplo, cuando el vehículo se balancea en los baches, resaltos, o curvas pronunciadas. Esto ha motivado en muchos trabajos la inclusión de mecanismos para medir o estimar en tiempo real la inclinación del vehículo respecto de la carretera (Nedevschi y otros, 2005; Suttorp y Bucher, 2006; Wang y Chen, 2006; Zhou y otros, 2006).

²traducido del inglés *clues*

Existen formas de relajar la hipótesis de mundo plano, eliminando la restricción de que el plano de la carretera y el del vehículo son el mismo, y aproximando la curvatura vertical por una recta, lo cual reduce considerablemente los errores en la estimación de la posición real de los objetos ([Guiducci, 1999](#)).

- **Colina-valle** (en inglés “*hill-and-dale*”): Este modelo emplea la aproximación de mundo plano para calcular la posición de la pareja de puntos (borde izquierdo y borde derecho del carril) más próximos al vehículo, a partir de los cuales se calcula el ancho del carril. Después, calcula la intersección de los haces de luz correspondientes a otras parejas de puntos con planos paralelos al original, pero a diferentes alturas, eligiendo en cada caso el plano que mantenga siempre el mismo ancho. Es decir, se obliga al plano de la carretera a elevarse o descender para mantener un ancho constante.
- **Sin peralte** (en inglés “*Zero-bank*”): Es un modelo propuesto por [DeMenthon \(1987\)](#), que considera la carretera como formada por una línea central, a modo de “espinas dorsales”, y una serie de segmentos horizontales que la cruzan perpendicularmente, dispuestos a intervalos regulares como de las traviesas de una vía férrea se tratase. La carretera queda formada por una sucesión de trapecios combados unidos entre sí. Años después, [Kanatani y Watanabe \(1990\)](#) proponen un modelo similar que reduce los trapecios a pequeños diferenciales, de modo que la carretera queda expresada en términos de un conjunto de ecuaciones diferenciales, y por lo tanto siguiendo una curva continua y diferenciable. Este modelo permite reconstruir la curvatura horizontal y vertical, pero tiene problemas de inestabilidad y sensibilidad al ruido, ya que debido a la necesaria integración numérica, los errores se van acumulando a lo largo del perfil de la carretera.

Otro de los modelos clásicos es el propuesto por [Dickmanns y Mysliwetz \(1992\)](#):

- **Espacio-temporal**: Considera que la carretera sigue el perfil de la curva de la clotoide, que es el utilizado en el diseño de las carreteras. Se denomina “espacio-temporal” ó “4D” porque incluye restricciones tanto en el dominio espacial como en el temporal (la cuarta dimensión). Es un modelo de variables de estado que combina la dinámica del vehículo con la geometría de la carretera. La geometría, a su vez, se puede separar en curvatura horizontal y vertical, ya que al ser éstas expresadas en términos de ecuaciones diferenciales, se pueden considerar independientes. Se trata de un modelo muy completo, pero de gran complejidad. Es sin embargo el más citado en la bibliografía, y constituye la base de muchos de los trabajos posteriores.

Existen, por otro lado, varias formas de modelar la curvatura horizontal y vertical de la carretera, como se verá en el apartado de enfoques descendentes.

A continuación se describen distintas técnicas utilizadas en detección de carriles, clasificadas en los dos bloques que ya se enunciaron en el apartado 2.2.3:

- Los métodos *ascendentes* o basados en características,
- y los métodos *descendentes* o basados en modelos.

2.5.1. Enfoques ascendentes

Estas técnicas suelen tener un fuerte preprocesado para extraer los objetos candidatos a bordes de carril. En una etapa posterior estos objetos se filtran y agrupan entre sí para formar objetos de más alto nivel.

Son frecuentes las técnicas que agrupan objetos basándose en relaciones geométricas:

En (Foresti y Regazzoni, 1999) se detectan líneas rectas y círculos mediante la transformada de Hough, en una imagen de bordes obtenidos con el filtro de Canny. Estos objetos, denominados primitivas, se unen mediante grafos que los relacionan entre sí en base a propiedades geométricas (simetría, paralelismo, convergencia, etc.), para formar objetos de más alto nivel (peatones, vehículos y marcas viales). A estos grafos se les aplica un modelo de campo aleatorio de Markov –“*Markov Random Field*” en inglés–, y mediante templado simulado –“*simulated annealing*” en inglés– se obtiene la mejor configuración.

En (Gonzalez y Ozguner, 2000), una vez segmentados los elementos pertenecientes a marcas viales, se les da una medida de confianza basada en propiedades relacionadas con su geometría. A continuación, se realiza una clasificación relacional de cada objeto mediante un árbol de decisión.

En (Broggi y otros, 1999) se aplica un pre-procesado basado en gradiente y operaciones morfológicas para extraer pequeños segmentos pertenecientes a marca vial, que se agrupan entre sí procurando seguir un modelo parabólico. Utiliza un filtrado temporal básico que consiste en limitar la región de análisis a una zona próxima a las marcas detectadas en la observación anterior.

En cambio, el sistema RALPH (Pomerleau y Jochem, 1996) utiliza una estrategia completamente distinta. El objetivo de este sistema es obtener simplemente la posición lateral del vehículo, de modo que hipotetiza cinco posibles modelos de curva y para cada uno de ellos desplaza las filas para corregir la curvatura. La imagen en la que se aprecian dos fuertes picos en la proyección sobre la horizontal corresponde a la hipótesis más acertada, y de ahí se extrae la curvatura de la carretera y la posición lateral comparando la posición del vehículo con un modelo previamente almacenado. El modelo se actualiza dinámicamente de forma automática, o por intervención del usuario.

También son muy utilizados los sistemas de votación, al estilo de la Transformada de Hough:

De la Escalera (1995) estudia ciertas ventanas colocadas en zonas donde se predice la existencia de marcas en base a las observaciones anteriores. En cada ventana se aplica la Transformada de Hough a los puntos de máximo gradiente, y uniendo los segmentos detectados se reconstruye la geometría de la carretera.

En Watanabe y Nishida (2005) se detectan las líneas con la transformada de Hough, partiendo de una imagen de bordes. Después, se exploran todas las parejas de líneas, eliminando, en caso de cruce, las líneas menos votadas. Finalmente se selecciona la pareja cuyo ancho sea más parecido a la detección en el fotograma anterior, y con esta observación se estiman los parámetros de un modelo recto de carril mediante el filtro de Kalman.

Otros sistemas buscan las marcas viales explorando la imagen secuencialmente, de abajo hacia arriba:

En (Taoka y otros, 2007) la imagen se divide en región *próxima* y región *lejana*. En la región próxima, se agrupan, mediante K-medias, los píxeles de marca vial en dos clases, una para el borde izquierdo y otra para el borde derecho del carril. A cada clase se le aplica la transformada de Hough, y de esa manera detecta los bordes del carril en la región próxima. En la región lejana, continua buscando las marcas viales mediante lo que el autor denomina “*step-walking*”, que consiste en partir del último punto detectado, y explorar una ventana con forma de semicírculo, colocada inmediatamente por encima de dicho punto, y dividida en subventanas que abarcan ángulos iguales. La subventana que contenga mayor número de píxeles de marca vial es la que indica el ángulo en que se debe seguir buscando la línea vial. Finalmente los puntos son sustituidos por una curva Bezier.

En (Cheng y otros, 2006), una vez segmentadas las marcas viales, se construye una lista de objetos a seguir, en función del tamaño y forma de los objetos –“*blobs*” en inglés–, y se descartan aquellos que no siguen el patrón de movimiento de la carretera. En la parte inferior de la imagen, mediante un sistema de votación similar a la transformada de Hough, se detectan los bordes izquierdo y derecho. Se recorre cada línea, y en el momento en que se detecta que hay un número significativo de píxeles de marca vial que no pertenecen a la línea, ésta se interrumpe y se fija un “punto de giro”. A partir de dicho punto se repite el proceso de votación, realizando un barrido angular para explorar todas las líneas que parten de él, y eligiendo la línea que contiene mayor número de píxeles de marca vial. En el caso de que el ancho de carril difiera significativamente del detectado en el fotograma anterior, se descarta la observación actual y se reemplaza por la última detección. El sistema avisa cuando no se ha encontrado el carril en varios fotogramas consecutivos.

Kim (2006) utiliza splines cúbicos como modelo de línea vial, que localiza en la imagen mediante una combinación del algoritmo RANSAC y el filtro de partículas. Mediante RANSAC, ayudado por el filtro de partículas para darle una coherencia temporal, se generan una serie de hipótesis de líneas viales que se puntúan en función del número de píxeles de marca vial que contienen. Se selecciona un pequeño número de hipótesis, se agrupan en función del “*lane probability score*”, que es una medida que penaliza las variaciones bruscas de ancho de carril, y se selecciona la pareja más probable.

La ventaja de estos métodos es que al no partir de un conocimiento a priori del mundo tan restrictivo como los métodos descendentes, permiten localizar todas las marcas viales presentes, y reconocer situaciones atípicas como carriles de ancho variable, bifurcaciones, carriles de entrada y salida, etc. La desventaja es que al cubrir tanta variedad de situaciones se hacen más vulnerables a las oclusiones y a la aparición de falsos positivos, debidos al ruido o la presencia en la imagen de otras estructuras de apariencia similar (ruido estructurado). Es por ello que se requieren complejas técnicas de validación de las observaciones para dar una cierta robustez al sistema. Por otra parte la ausencia de un modelo complica la aplicación de técnicas estadísticas de seguimiento –“*tracking*” en inglés–.

Con todo, y como veremos en siguientes apartados algunas de estas técnicas han superado con bastante éxito pruebas exhaustivas de conducción automática (Broggi y otros, 1999; Pomerleau y Jochem, 1996).

2.5.2. Enfoques descendentes

Los enfoques descendentes o basados en modelos parten de un modelo deformable predefinido, que reúne el conocimiento “a priori” del objeto que se quiere localizar, y que se ajusta a las observaciones mediante alguna técnica. Cuando el objeto que se busca es la carretera, el modelo debe dar una representación precisa de la misma, ser robusto frente a oclusiones, calidad de las marcas y condiciones ambientales, y ser eficiente computacionalmente.

Los distintos modelos utilizados se pueden clasificar atendiendo al número de líneas viales que ajustan –generalmente son dos, aunque algunas propuestas admiten más–, al tipo de perfil de la carretera –línea recta, arco de circunferencia, parábola, cúbica, clotoide, “spline”, “snake”–, o a la consideración o no de otros aspectos como la curvatura vertical, la posibilidad de variar en anchura, el peralte o la coherencia temporal a lo largo de una secuencia de imágenes.

Atendiendo al perfil de la carretera, los bordes del carril se pueden modelar como:

- **Líneas rectas** (Foda y Dawoud, 2001; Apostoloff y Zelinsky, 2003; Voisin y otros, 2005): es el modelo más sencillo. Resulta fácil y rápido de ajustar, pero no da información de la curvatura de la carretera.

- **Arcos de circunferencia** (Guiducci, 1999) o **Parábolas** (Kreucher y Lakshmanan, 1999; Park y otros, 2003; Li y otros, 2004; Watanabe y Nishida, 2005; McCall y Trivedi, 2006; Wang y Chen, 2006; Zhou y otros, 2006): Son una buena aproximación en carreteras de curvatura suave, como las autopistas. En carreteras de menor categoría, con curvas bruscas o en “S”, la parábola sólo puede aproximar el entorno cercano, ya que no permite puntos de inflexión.
- **Cúbicas** (Chapuis y otros, 2002; Southall y Taylor, 2001; Nedevschi y otros, 2005): Permiten aproximar muy bien el perfil de la carretera, pero al tener un mayor número de grados de libertad son también más difíciles de ajustar debido al ruido. Se suelen utilizar como aproximación a las clotoides (Risack y otros, 1998).
- **Clotoides** (Dickmanns y Mysliwetz, 1992): Se definen como curvas cuyo ritmo de variación de curvatura es constante. Este perfil se utiliza en el trazado en planta de las carreteras (BOE, 2000, cap. 4), en concreto como curva de transición entre los otros dos tipos de elementos que podemos encontrar: las rectas y las curvas circulares. Si bien, las rectas pueden considerarse como casos particulares de clotoides, esto no sucede con las curvas circulares. A pesar de ello, este es el modelo que mejor se aproxima a la realidad en la mayoría de los casos, pero al estar expresado en coordenadas paramétricas la matemática posterior se complica, por lo que con frecuencia se aproxima por una cúbica.
- **Contornos deformables**: Por ejemplo, *splines* (Wang y otros, 2000) o *snakes* (Yuille y Coughlan, 2000; Wang y otros, 2004; Kim, 2006). Pueden aproximar muy bien cualquier perfil, y aunque los puntos de control varían menos entre fotograma y fotograma que los parámetros de los modelos anteriores, son más difíciles de seguir en una secuencia de vídeo mediante técnicas como el filtro de Kalman o el de partículas, ya que su modelado es más complejo.
- **Modelos mixtos** (zona lineal y zona parabólica) (Jung y Kelber, 2005; Lei y otros, 2006): Ajustan un modelo lineal en el entorno cercano, y un modelo parabólico en el lejano. De esta forma aprovechan la sencillez del modelo lineal, que una vez ajustado limita enormemente la zona de búsqueda del modelo parabólico.

De la misma forma, en relación a la curvatura vertical, podemos distinguir las siguientes aproximaciones:

- **Mundo plano**: La curvatura vertical se considera despreciable (figura 2.7a), o en el mejor de los casos se hace un seguimiento de la línea de horizonte o punto de fuga, para actualizar el plano en el entorno cercano (figura 2.7b).

- **Plano inclinado** (Guiducci, 1999; Jeong y Nedevschi, 2005): Se diferencia del caso anterior en que, en lugar de modificar la inclinación de un plano que parte de la base del vehículo, se ajusta el perfil real a una recta mediante regresión lineal (figura 2.7c).
- **Parábola** (Dickmanns y Mysliwetz, 1992; Chapuis y otros, 2002; Nedevschi y otros, 2005) Como el caso anterior, pero ajustando una parábola (figura 2.7d). Éste es el perfil que se utiliza en el trazado en alzado de las carreteras, como acuerdo vertical entre dos rasantes (BOE, 2000, cap. 5).

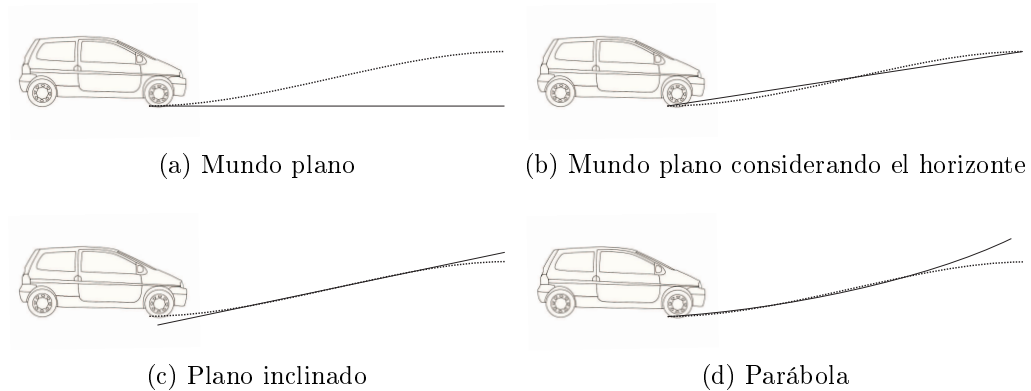


Figura 2.7: Modelos de curvatura vertical.

El ajuste del modelo se puede hacer sobre el plano de la imagen o sobre el de la carretera. Conociendo la altura e inclinación de la cámara y suponiendo que la carretera es plana (Broggi y otros, 1999; Vlacic y otros, 2001; McCall y Trivedi, 2006) se puede obtener una imagen de la carretera a vista de pájaro. De esta forma el ajuste se puede hacer directamente en coordenadas del mundo, en lugar de la imagen.

Para eliminar el ruido y las falsas detecciones, se suele hacer un seguimiento temporal de las observaciones. Para esto hay básicamente dos alternativas (Ieng y otros, 2003):

- El *filtro de Kalman* (Bar-Shalom y Fortmann, 1988), ampliamente utilizado en el mundo de la robótica. Su principal dificultad reside en el modelado del sistema y la inicialización del algoritmo, de la que depende en gran medida su convergencia.
- El *filtro de partículas* (Doucet y otros, 2001) es una técnica de uso más reciente. Supone mayor coste computacional que el filtro de Kalman, pero permite trabajar con distribuciones multimodales, y tiene más facilidad para recuperarse en caso de pérdida del objeto a seguir.

En (Dickmanns y Mysliwetz, 1992) se define el modelo *espacio-temporal* o $4D$, haciendo alusión a la cuarta dimensión que corresponde al tiempo. Este modelo contempla la curvatura horizontal y vertical de la carretera. Ambas curvaturas se suponen independientes, y se modelan con clotoideas. La curvatura vertical se calcula asumiendo que el ancho de carril es constante. Además de los parámetros geométricos de la carretera y de calibración de la cámara, este modelo incluye las ecuaciones diferenciales que relacionan movimiento del vehículo con los desplazamientos espaciales. El modelo es muy completo, pero presenta una alta sensibilidad al ruido dada la gran cantidad de grados de libertad. Este problema puede paliarse mediante la integración de la información de varios sensores. No obstante, el modelo espacio-temporal es muy citado, y con frecuencia se utilizan modelos similares simplificados.

El trabajo de Campbell y Thomas (1993) es de los pocos que modelan directamente las líneas discontinuas. Trabaja en el plano de la carretera, y utiliza un modelo marca vial con seis parámetros, de los cuales tres corresponden a la forma parabólica, y otros tres al período, ciclo de trabajo y fase de una onda. El modelo se actualiza mediante mínimos cuadrados no lineales. Éste modelo es interesante ya que permite modelar cualquier tipo de marca vial, pero su inicialización previa al seguimiento es muy costosa computacionalmente.

Tanto en LOIS (Kluge y Lakshmanan, 1995) como en LANA (Kreucher y Lakshmanan, 1999) se parte de la aproximación de “mundo plano”, y se asume que las marcas son arcos circulares sobre el plano de la carretera, y que en el plano de la imagen se pueden aproximar por parábolas. Trabajando las ecuaciones llegan a un modelo de parábola en el que sólo uno de los tres parámetros varía al pasar de un borde al otro. Por tanto con cuatro parámetros se llega a definir un modelo de carretera delimitado por dos parábolas. Después se ajusta el modelo a la imagen con un enfoque bayesiano.

En (Risack y otros, 1998) se utiliza un modelo espacio-temporal, pero simplificando la dinámica del vehículo para utilizar solo información de la velocidad. El vector de estados incluye las curvaturas horizontal y vertical (supuestas independientes), el ancho del carril, y la orientación del vehículo y de la cámara. Este vector se ajusta a las observaciones mediante mínimos cuadrados, y los resultados se validan en base a varias reglas geométricas. Es interesante en este trabajo la detección de carriles adyacentes, asumiendo que algunos de sus parámetros son idénticos a los del carril principal, y la clasificación de marcas en continua/discontinua. Para ello analiza las discontinuidades en los bordes detectados para cada carril, y cuando se encuentra una cierta separación se califica la línea como discontinua. Obviamente cualquier oclusión de una línea continua dará lugar a una clasificación errónea. Las oclusiones también pueden afectar bastante a la etapa de detección, ya que sólo se analizan una serie de líneas horizontales de búsqueda, lo que hace al algoritmo sensible al ruido, y susceptible de perder observaciones si las líneas no están bien marcadas, o hay parches de asfalto o sombras.

En (Guiducci, 1999) se considera la carretera como un arco de circunferencia, que se proyecta sobre la cámara como una hipérbola. Las hipérbolas las ajusta mediante mínimos cuadrados, y se siguen en la secuencia mediante el filtro de Kalman. Relaja la hipótesis de mundo plano considerando que el plano de la carretera se ajusta al perfil vertical de la misma en un punto central en la longitud de carretera analizada. Según los resultados obtenidos en el vehículo MOBLAB, esta aproximación sólo provoca errores de más de un píxel en el 2 % de los fotogramas analizados.

En (Wang y otros, 2000) se ajustan los bordes izquierdo y derecho del carril con splines, cada una de ellas con tres puntos de control, uno inicial, otro final, y uno intermedio. La búsqueda se hace en la imagen del gradiente de intensidades. El primer punto se coloca en la parte inferior de la imagen. El último sobre el horizonte, cuya posición se deduce calculando la intersección de dos líneas ajustadas mediante la Transformada de Hough aplicada una imagen de bordes. La búsqueda de los puntos de control no es muy robusta, especialmente ante oclusiones. Requiere líneas preferiblemente continuas, bien marcadas y libres de obstáculos para funcionar correctamente.

En (Southall y Taylor, 2001) se parte de la hipótesis de carretera plana y se modela su curvatura horizontal con una cúbica. Se define un modelo de seis variables de estado, que incluyen el desplazamiento lateral, la inclinación del vehículo respecto al carril, la curvatura y anchura del carril, y la inclinación de la CCD. Es interesante la actualización dinámica de la inclinación de la cámara, haciendo un seguimiento de la línea del horizonte mediante correlación. El seguimiento de la carretera se hace con el algoritmo CONDENSATION (Isard y Blake, 1998). Aparentemente el algoritmo tiene problemas con las entradas y salidas de curvas, y como en el caso anterior aún no ha sido muy probado.

En el sistema VELAC (Chapuis y otros, 2002) se utiliza un modelo con curvaturas horizontal y vertical parabólicas en una carretera de anchura variable y sin peralte. La detección de marcas se hace definiendo intervalos de confianza para cada una de ellas, que va reduciendo mediante el análisis de pequeñas regiones en la imagen en las que busca puntos de alto gradiente. El ruido se elimina mediante el filtro de Kalman. El modelo teórico es interesante, aunque la discusión de resultados en este artículo es escasa debido a la ausencia de pruebas más exhaustivas.

En (Apostoloff y Zelinsky, 2003) se ajusta un modelo consistente en dos líneas rectas. Lo interesante de su trabajo es, aparte de la etapa de extracción de características que ya se comentó, la etapa de seguimiento, que utiliza el “*Distillation Algorithm*” que combina el filtro de partículas y la fusión de múltiples “estímulos”³. Este algoritmo tiene el inconveniente de que al seleccionar una “partícula” concreta cada vez, el seguimiento no es suave, sino a saltos. Con el modelo de línea recta estima el ancho de la carretera, que puede variar, y la posición y orientación del vehículo dentro de ella.

³traducido del inglés “cue”

En [Park y otros \(2003\)](#), se considera que la carretera sigue un modelo parabólico, y que se proyecta en la imagen como unas curvas cónicas que el autor denomina “*Lane Curve Function*” (LCF). En el entorno cercano aproxima las líneas viales por una recta. Después, define un índice de curvatura que calcula mediante un sistema de votación. Para cada índice de curvatura, calcula un histograma de similitud entre orientación de bordes real y esperada (es decir, cuántos píxeles cuya orientación se corresponde con esa curvatura). La curvatura real corresponde al histograma de mayor área. En el artículo sólo se mencionan pruebas con secuencias fuera de línea (“*off-line*”).

En [Nedevschi y otros, 2005\)](#) se modela el perfil horizontal y vertical de la carretera como clotoides aproximadas por una curva cúbica y una parábola respectivamente, apoyándose en un sistema estéreo. En el pre-procesado extrae bordes eliminando zonas cuyos bordes no tienen una dirección predominante. Después se obtiene la posición 3D de los puntos de bordes aplicando estéreo, y con esta información extrae el perfil vertical de la carretera en dos pasos. Primero calcula con la Transformada de Hough el ángulo de cabeceo del vehículo utilizando puntos próximos, y a continuación, suponiendo dicho ángulo conocido, calcula el segundo parámetro de la parábola también con Hough. Una vez estimado el perfil vertical, elimina los bordes que no pertenecen al plano de la carretera, y crea una zona de búsqueda en torno a la estimación de carril dada por el filtro de Kalman, actualiza los parámetros de la curvatura horizontal y vertical ajustando las clotoides a los puntos englobados en la zona de búsqueda.

En [Jung y Kelber, 2005\)](#) se considera que la carretera sigue un modelo lineal en el entorno próximo (hasta unos 30-40m del vehículo), mientras que en el entorno lejano el modelo es parabólico, en este caso en el sistema de coordenadas de la imagen. Imponiendo a ambas curvas la condición de continuidad y derivabilidad, obtiene un modelo *lineal-parabólico* de tres parámetros para cada borde. Una vez segmentados los bordes de la carretera, calcula el modelo lineal mediante la transformada de Hough, y el modelo parabólico por minimización de error cuadrático. No hace seguimiento como tal, pero limita la zona de búsqueda a las proximidades del carril detectado en el fotograma anterior. En las pruebas de este método la detección oscilaba en el entorno lejano cuando los bordes no estaban bien marcados. En general, el algoritmo se distrae cuando hay fuertes bordes que no pertenecen a marcas viales, como los provocados por sombras o por el contraste entre la calzada y la vegetación.

El sistema VioLET ([McCall y Trivedi, 2006\)](#) considera la carretera como una parábola sobre un mundo plano. Procesa la imagen en perspectiva inversa, y la detección del perfil la hace en dos pasos. Primero ajusta líneas rectas con la transformada de Hough, y después estima la curvatura mediante correlación (“*template matching*”). En la imagen en perspectiva inversa, selecciona un recuadro sobre una zona próxima al vehículo y por donde se estima que el carril pasa. Esa región se va ajustando en varios puntos equidistantes entre

sí, a lo largo de la carretera, de forma que el mejor ajuste da una estimación de la posición del centro del carril. A continuación, todos esos puntos se ajustan a una parábola. Obviamente siempre que haya discontinuidades en la apariencia de la carretera, a medida que nos alejamos del vehículo, por ejemplo, en presencia de vehículos o sombras. En estos casos, utiliza la información inercial del vehículo para estimar la posición del carril. Con el filtro de Kalman actualiza el ancho y dirección de la carretera. Este sistema ha sido probado en el vehículo LISA-Q, a lo largo de una ruta de 65km por San Diego (California). Para validar la detección del carril, una cámara enfocada al suelo proporciona información de distancia lateral. En la evaluación concluye que el “template-matching” falla en carreteras con textura pobre y oclusiones. El algoritmo se comporta mejor de noche, porque hay mayor contraste, y al amanecer, porque la niebla matinal reduce las sombras. Sin embargo, los cambios de iluminación requieren variaciones en los umbrales de detección de marcas viales.

En (Zhou y otros, 2006) también se considera la carretera como una parábola, que se proyecta sobre la imagen como una hipérbola. Lo característico de este trabajo es la aplicación de la *búsqueda tabú* al ajuste de modelo en la imagen. El seguimiento del carril, se hace con un filtro de partículas. Este sistema ha sido probado con secuencias de imágenes grabadas a 25 fotogramas por segundo. En el artículo se afirma que el sistema es más robusto frente a discontinuidades en las líneas viales que el propuesto en (Guiducci, 1999), y que la búsqueda tabú funciona mejor que el algoritmo METROPOLIS (Kluge y Lakshmanan, 1995). También se sugiere que, en presencia de vehículos, funcionará mejor una combinación de algoritmos de detección (búsqueda tabú) y seguimiento (filtro partículas).

Como conclusión, puede decirse que en general estos enfoques, al aprovechar todo el conocimiento a priori que se tiene del entorno y que queda incluido en la definición del modelo de carretera, son más robustos que los ascendentes y presentan menor número de falsos positivos. El uso de modelos deformables de carretera, combinado con modelos de la dinámica del vehículo facilita el seguimiento temporal de la carretera. Si además se incluye información de sensores inerciales del vehículo, se pueden tolerar breves pérdidas de la visibilidad de los carriles.

Su mayor inconveniente es la incapacidad de tratar con situaciones que se salgan de ese conocimiento a priori previamente establecido. No puede ser reconocida ninguna situación que no haya sido incorporada en el modelo.

Por ejemplo, la mayoría de los modelos sólo consideran la existencia de un carril. Existen algunos trabajos que tratan también de localizar carriles adyacentes. En (Risack y otros, 1998) se exploran dos marcas más, una a la derecha y otra a la izquierda, suponiendo que comparten ciertos parámetros con las marcas del carril ya detectado. En (Aufrère y otros, 2001) también se contempla la existencia de carriles adyacentes. Se crea una tabla de probabilidad de existencia de otros carriles, y cuando la probabilidad es alta, se

activa en el siguiente fotograma la búsqueda de otra marca más, siguiendo la misma técnica que para el carril central. En (Labayrade y otros, 2006) se propone un algoritmo que busca varios modelos en una misma imagen. Una vez localizado un modelo, se “eliminan” los píxeles de uno de los bordes, y se busca un nuevo modelo. El proceso se realiza igual con ambos bordes. Si uno de estos nuevos modelos se localiza con una cierta probabilidad, se incluye en la tarea de seguimiento. El objetivo es poder reconocer “singularidades” como carriles de entrada/salida, o líneas dobles de separación de carriles. Sin embargo, son aún escasos los trabajos de investigación que van en esta línea.

En este sentido, una de las aportaciones de esta tesis consiste en la definición de un modelo capaz de reconocer situaciones que actualmente han recibido un escaso tratamiento por parte de los trabajos encontrados en la bibliografía, como es el caso de la clasificación automática del tipo de líneas viales, y el reconocimiento automático del número de carriles. Esta información permitirá en el futuro la identificación del tipo de carretera por la que circula el vehículo.

En relación a las técnicas empleadas, se ha detectado una amplia variabilidad de métodos de ajuste de modelos, que contrasta con el reducido número de técnicas de seguimiento, prácticamente limitado a los clásicos filtros de Kalman y de partículas. Dado que no existe actualmente una base de datos común de imágenes de carreteras etiquetadas, no es posible establecer una comparativa entre todas las técnicas utilizadas, en relación a su eficiencia y eficacia. Se aprecia por tanto una interesante línea de investigación consistente en la definición de una metodología de evaluación sistemática que permita comparar estos sistemas entre sí.

Respecto a los sistemas de visión, existen técnicas que permiten reconstruir la curvatura horizontal y vertical de la carretera con un sistema monocular (Chapuis y otros, 2002; Glaser y otros, 2006), generalmente basándose en la hipótesis de ancho constante y peralte nulo. Un sistema estéreo permite también identificar la curvatura horizontal y vertical, sin necesidad de recurrir a dichas restricciones (Nedevschi y otros, 2005; Sappa y otros, 2006). Apenas hay trabajos que consideren el peralte de la carretera, pero sí existen técnicas que reconocen el ángulo de alabeo de la cámara utilizando un sistema estéreo (Labayrade y Aubert, 2003; Nedevschi y otros, 2005).

Otros inconvenientes de los enfoques descendentes son la dificultad de elegir y mantener el modelo apropiado, ya que diferentes entornos pueden requerir diferentes modelos, y la alta complejidad computacional que requiere el ajuste de modelos complejos. A medida que se añaden grados de libertad al modelo, éste se hace más inestable y sensible a pequeñas variaciones, como por ejemplo, la del ángulo de inclinación de la cámara debido a los baches y a la suspensión.

2.6. Proyectos clásicos

Se detallan varios proyectos relevantes por que han sido expuestos a test exhaustivos, y por que ofrecen una información más detallada de su funcionamiento. Por otra parte, los cuatro proyectos aquí descritos siguen enfoques muy diferentes y sirven para ilustrar la variedad de técnicas de visión que se pueden aplicar a la detección de carreteras. La mayoría de estos proyectos están diseñados para la conducción automática. Los dos primeros, ALVINN ([Pomerleau y Jochem, 1996](#)) y RALPH, no hacen una detección explícita de las marcas de la carretera. Todo el procesado de ALVINN se basa en redes neuronales, mientras que RALPH, siguiendo un enfoque ascendente, busca características de la carretera que sean paralelas a su eje, como las marcas viales o huellas de neumáticos. El vehículo ARGO sigue también un enfoque ascendente, pero en este caso sí hace detección explícita de las líneas viales. Por el contrario, el vehículo UTA II de la DaymleChrysler es un ejemplo de un enfoque descendente, en el que la detección de carriles sigue el esquema muy similar al propuesto en ([Dickmanns y Mysliwetz, 1992](#)), con cuyo autor colaboró durante largo tiempo.

2.6.1. ALVINN (Carnegie Mellon University)

El proyecto ALVINN (“Autonomous Land Vehicle In a Neural Network”) ([Pomerleau, 1993](#)) utiliza una sola red de neuronas para las funciones de percepción y control. El hecho de dejar el preprocesado de la imagen a la red surge de la intención de poder aplicar el sistema a diversos modelos de carretera, sin necesidad de determinar previamente qué características son relevantes en cada uno, y dejar que sea la propia red quien lo decida. Escoge un perceptrón multicapa basándose en los criterios de que la red debe poder ser entrenada de forma supervisada para obtener las salidas deseadas, y que las dependencias temporales no son relevantes para el análisis. Estos criterios excluyen a los mapas de Kohonen y las redes recurrentes.

Arquitectura de la red

La arquitectura de la red se muestra en la figura [2.8a](#). La entrada consiste en una matriz de neuronas de 30x32 píxeles –a modo de retina–, la capa oculta tiene cuatro neuronas, y la salida consta de treinta, que representan una discretización del giro del volante (cada neurona representa un ángulo). El giro máximo corresponde a una curva de 20m de radio.

En la entrada se añade una capa más para procesar el color, de forma que a cada neurona de la retina le llega la salida de otras tres correspondientes a las tres componentes de color (RGB). Para simplificar el aprendizaje se da el mismo peso a todas las neuronas de la misma componente de color, de forma que sólo hay que aprender tres pesos en esta capa.

En la última capa (figura 2.8b), las neuronas se modifican para generar una salida que pueda recorrer el intervalo real $[-1,1]$, en lugar de tener una salida binaria. Esta capa se entrena para generar una gaussiana centrada en el ángulo de giro del volante. Como salida de la red se toma el centro de masas de las activaciones de las neuronas. Así se incrementa la precisión de la salida, al eliminar la limitación impuesta por la discretización del intervalo de giro en 30 valores.

Entrenamiento

El entrenamiento de ALVINN destaca por su originalidad. El algoritmo utilizado es el de retropropagación con ligeras modificaciones. Lo particular del entrenamiento es la forma de generar los patrones de ejemplo. Se pretende dotar al sistema de un aprendizaje “al vuelo”, es decir, que pueda aprender observando la conducción de un humano. Pero aquí surge el problema de que el humano siempre conducirá centrado en el carril, de tal forma que la red nunca se verá expuesta a situaciones anómalas de las que deba recuperarse, ya que si el vehículo se saliese de la carretera, la red trataría de aprender también ese movimiento. Para compensar esto, a partir de las imágenes captadas se generan imágenes virtuales desplazadas ligeramente hacia los lados (ver figura 2.8c), corrigiendo la perspectiva. La dirección correcta será el desplazamiento introducido en la imagen. Otro problema es que el conjunto de ejemplos está ordenado, lo que facilita el “olvido” de lo aprendido tiempo atrás. Para paliar este efecto, se mantiene un pequeño “buffer” con ejemplos de situaciones pasadas, que se le introducen a la red de vez en cuando para refrescar.

Resultados

ALVINN ha conducido el vehículo NavLab sin intervención humana durante un recorrido de 22 millas (35,4 km). El sistema procesa 15 imágenes por segundo (en su momento fue el más rápido), permitiendo conducir a 55 millas/h (88 km/h). En autopista llega a alcanzar las 90 millas/h (145 km/h).

ALVINN busca características de la imagen de carretera en las que fijarse, pero no está asegurado que estas características sean las adecuadas. Si hay presente un ruido estructurado y continuo en el tiempo, la red acabará aprendiéndolo. Un ejemplo de esto sucedió al circular por un camino de tierra con una zanja a la derecha. El análisis de la capa oculta reveló que la red aprendió a reconocer débilmente la carretera y se fijó bastante en la zanja de la derecha. Al cambiar el sentido del recorrido el sistema daba bandazos pues la zanja estaba ahora a la izquierda.

Otras fuentes de error son: la dificultad que tiene el perceptrón para recordar situaciones poco comunes que hay en los ejemplos, y, debido al entrenamiento basado en la conducción humana, otras dos características que hacen difícil mantener el conjunto de ejemplos: el limitado conjunto de

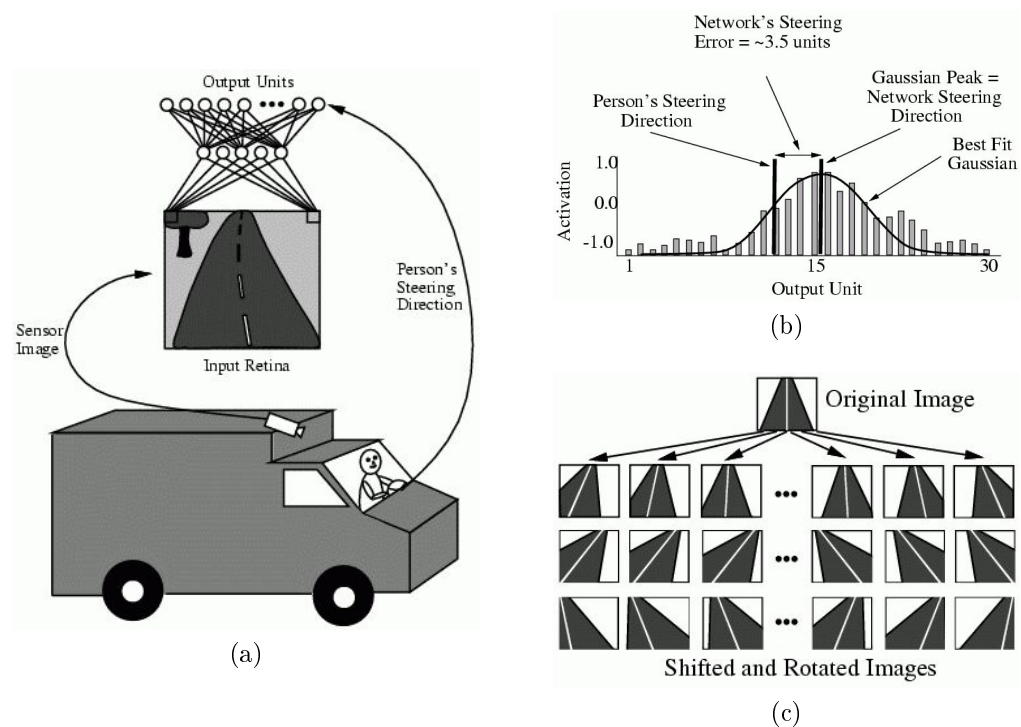


Figura 2.8: ALVINN, en (Pomerleau, 1993).

situaciones y las correlaciones temporales (que antes se mencionaban como no relevantes, pero que a la luz de los resultados sí lo son). También debido al entrenamiento ALVINN sufre de pérdida de memoria, ya que los últimos ejemplos siempre tendrán más peso que los anteriores. Otro problema son las respuestas impredecibles de ALVINN frente al ruido estructurado.

2.6.2. RALPH (Carnegie Mellon University)

RALPH (*“Rapidly Adapting Lateral Position Handler”*) (Pomerleau y Jochem, 1996) es uno de los sistemas que ha sido sometido a un test intensivo, denominado *“No hands across America”* (Jochem y Pomerleau, 1995), en el que se realizó un viaje de 2.849 millas (4.585 km), de costa a costa de Estados Unidos. RALPH logró conducir un 98,2% del trayecto en automático.

RALPH ha sido desarrollado conjuntamente por la Carnegie Mellon University (CMU) y AssistWare Technology, Inc. (*“spin-off”* de la CMU). Este sistema ha sido diseñado para la conducción automática en autopistas y para la prevención de salidas no intencionadas de carril. Se puede configurar para dar avisos de salida de carril solamente, o para controlar el giro del volante de forma autónoma.

La determinación del giro del volante se descompone en tres pasos:

- 1) **Muestreo de la imagen:** El sistema muestrea únicamente una región trapecoidal de la imagen, que se corresponde con la proyección sobre

la cámara de un rectángulo de carretera en coordenadas del mundo, situado a 20m por delante del vehículo, de unos 7m de anchura y 50m de longitud. Las distancias son aproximadas, ya que el tamaño de esta ventana trapezoidal depende de la velocidad del vehículo (se agranda si la velocidad aumenta). Al muestrear la imagen, se crea otra de baja resolución (30x32 píxeles), en la que se ha deshecho la transformación de perspectiva, mostrando así una imagen a vista de pájaro (figura 2.9).

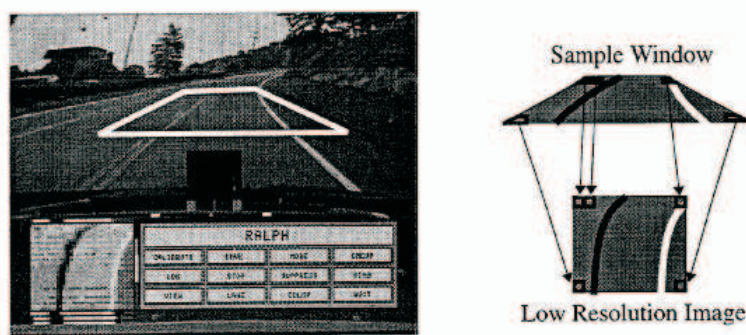


Figura 2.9: Muestreo de la imagen, con inversión de perspectiva, en (Pomerleau y Jochem, 1996).

- 2) **Determinación de la curvatura de la carretera:** En esta fase RALPH genera una serie de hipótesis de curvatura (figura 2.10a). Para cada una de ellas, RALPH desplaza las filas de la imagen como si quisiese rectificar la curva, según la curvatura que corresponde a la hipótesis. A continuación suma verticalmente las intensidades, para crear una proyección de la imagen sobre la horizontal. Así se crea un perfil de intensidades en el cual, si las líneas de la carretera se han rectificado, aparecerán fuertes picos, y en el caso de que las líneas de la carretera mantengan cierta curvatura, se apreciarán variaciones más suaves. RALPH selecciona de entre todas las hipótesis de curvatura, aquella que genera las variaciones más abruptas en la proyección sobre la horizontal (figura 2.10b).
- 3) **Cálculo de la posición lateral del vehículo:** Para este cálculo, RALPH mantiene un modelo que corresponde al perfil horizontal generado durante la determinación de la curvatura, cuando el vehículo se encuentra centrado en su carril. El cálculo de la posición lateral consiste simplemente en desplazar horizontalmente el perfil actual y correlarlo con el modelo almacenado. El punto de máxima correlación determina el mejor ajuste, y el desplazamiento horizontal del perfil se corresponde directamente con el del vehículo.

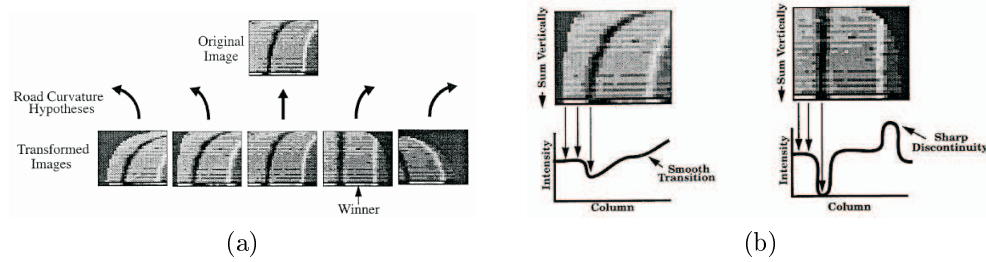


Figura 2.10: Estimación de la curvatura, en (Pomerleau y Jochem, 1996).

Actualización del modelo

Para adaptarse a condiciones cambiantes de entorno y de carretera, RALPH actualiza el modelo de perfil horizontal de cuatro posibles maneras:

1. El conductor sitúa al vehículo en el centro del carril y presiona un botón. Inmediatamente se almacena el modelo obtenido a partir de fotograma actual para ser utilizado en los fotogramas siguientes.
2. El sistema selecciona un modelo de una biblioteca de modelos previamente almacenados. El modelo seleccionado será el que mejor correlación tenga con el perfil creado a partir del fotograma actual.
3. El modelo se hace evolucionar añadiéndole un pequeño porcentaje del perfil actual. De esta forma el modelo se adapta gradualmente a ligeros cambios de iluminación, como los debidos al movimiento del Sol.
4. RALPH analiza continuamente una región lejana de carretera (aproximadamente de 70 a 100m), para la cual se crea un perfil asumiendo que la curvatura es la misma que para el entorno cercano. Si se produce un cambio brusco de entorno durante la conducción, RALPH determina que el modelo actual no es válido debido a un bajo índice de correlación, y sustituye el modelo antiguo con el obtenido para el entorno lejano.

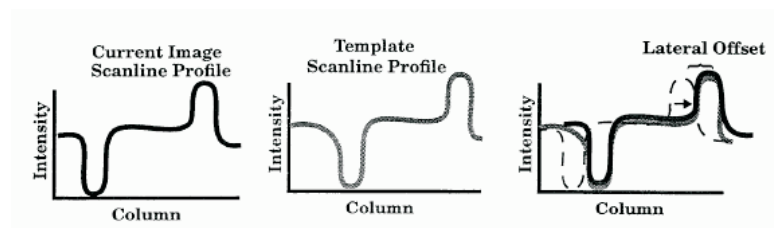


Figura 2.11: Estimación de la distancia lateral, en (Pomerleau y Jochem, 1996).

Evaluación del rendimiento

La gran ventaja de este sistema es su robustez. Debido a su peculiar etapa de percepción, RALPH es capaz de fijarse y seguir cualquier característica que sea paralela a la carretera. El sistema se pierde solamente cuando no hay ninguna característica que pueda seguir.

El inconveniente es la simplicidad del modelo de carretera y de sus capacidades perceptivas. Solamente es capaz de determinar de la carretera la distancia lateral y un valor aproximado de la curvatura horizontal.

2.6.3. ARGO (Universidad de Pavia)

El proyecto ARGO (Broggi y otros, 1999), desarrollado por las universidades de Parma y Pavia, tiene como principal objetivo el desarrollo de un sistema de seguridad activa con la capacidad de actuar también como piloto automático de un vehículo convencional. El vehículo ARGO está equipado con un sistema de visión estéreo y PCs convencionales.

Aquí describiremos la parte de detección de carreteras del sistema GOLD (en inglés, “*Generic Obstacle and Lane Detection*”). El sistema sigue el esquema descrito en el estado del arte, compuesto por una fase de extracción de características (bajo nivel), y posterior interpretación de las características (alto nivel) siguiendo un enfoque ascendente.

Procesamiento de bajo nivel

El primer paso es realizar la transformación de perspectiva inversa para obtener una imagen de la carretera a vista de pájaro, bajo la hipótesis de mundo plano. En esta imagen las marcas son bandas brillantes sobre fondo oscuro. Para extraer estas marcas, se compara la intensidad de un píxel $i(x, y)$ con la de sus vecinos izquierdo y derecho a una cierta distancia m en horizontal: $i(x, y - m)$ e $i(x, y + m)$, siendo $m \geq 1$.

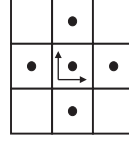
Se genera una imagen cuyos nuevos valores $r(x, y)$ codifican la presencia de marcas viales, siguiendo la expresión:

$$r(x, y) = \begin{cases} d_{+m}(x, y) + d_{-m}(x, y), & \text{si } d_{+m}(x, y) > 0 \wedge d_{-m}(x, y) > 0 \\ 0, & \text{en otro caso} \end{cases} \quad (2.1)$$

donde:

$$\begin{cases} d_{+m}(x, y) = i(x, y) - i(x, y + m) \\ d_{-m}(x, y) = i(x, y) - i(x, y - m) \end{cases} \quad (2.2)$$

Después, la imagen se realiza mediante un filtro no lineal denominado “*geodesic morphological dilation*” con el siguiente elemento estructural:



donde $c(x, y) = \begin{cases} 1, & \text{si } i(x, y) \neq 0 \\ 0, & \text{en otro caso} \end{cases}$ es la imagen de control.

La transformación consiste en sustituir el píxel central $i(x, y)$ por la máxima intensidad encontrada en el elemento de control, siempre que $c(x, y) = 1$.

A continuación se realiza una binarización con un umbral adaptativo:

$$t(x, y) = \begin{cases} 1, & \text{si } e(x, y) \geq \frac{m(x, y)}{k} \\ 0, & \text{en otro caso} \end{cases} \quad (2.3)$$

donde $e(x, y)$ representa la imagen mejorada, $m(x, y)$ el valor máximo calculado en una vecindad $c \times c$, y k es una constante.

Procesamiento de alto nivel

En el siguiente paso la imagen se explora para construir cadenas de píxeles, que posteriormente se ajustan con *poli-líneas*.

Primero se seleccionan aquellas *poli-líneas* próximas a las líneas viales detectadas en el fotograma anterior. Luego se concatenan unas con otras en base a unas reglas geométricas de cercanía en distancia y orientación. Después, las nuevas *poli-líneas* se evalúan para asignarles una puntuación. Aquellas que son cortas, que no cruzan la imagen de arriba abajo, o que tienen grandes variaciones angulares recibirán una penalización. Finalmente se selecciona la de mayor puntuación, y si no es suficientemente larga, se puede prolongar siguiendo una aproximación lineal o parabólica.

A continuación, la línea detectada se contrasta con un modelo interno de carretera consistente en dos líneas paralelas colocadas en una posición concreta. Este modelo se inicializa al comienzo de todo el proceso, y puede sufrir ligeras modificaciones durante la conducción para adaptarse a nuevas condiciones de carretera, gracias a un proceso de aprendizaje que se ejecuta en segundo plano.

Evaluación del sistema

El sistema fue probado en un test denominado “*MileMiglia in Automatico Tour*” (Conte y otros, 1998), que tuvo lugar del 1 al 6 de junio de 1998, y consistió en un recorrido de 2.000 km de conducción automática por autopistas italianas. El sistema demostró ser robusto ante la presencia de otras marcas viales, guarda-raíles, intersecciones y tráfico intenso. Esta prueba permitió extraer una serie de conclusiones que han motivado varias mejoras posteriores, y que presentamos a continuación:

- El uso de cámaras CCD, de bajo coste, con un lento control automático de ganancia, da lugar a una degradación de la imagen ante cambios bruscos de iluminación, como por ejemplo, en las entradas y salidas de los túneles. La cámara se satura por completo y durante un pequeño período de tiempo el procesado de la imagen es imposible.
- En los casos en que la aproximación de carretera plana no se cumple, las líneas que aparecen en la imagen de perspectiva inversa pierden su paralelismo, y como consecuencia, se produce una pequeña desviación en la estimación de la distancia lateral, aunque no lo suficientemente grande como para provocar la salida del carril.
- Durante la conducción, el sistema se fija generalmente en el borde derecho del carril, ya que como se ha comentado previamente se selecciona sólo una poli-línea, y el borde derecho, al ser continuo, da una observación más fiable. El problema surge cuando se presenta un carril de incorporación o salida, de modo que aparecen dos líneas a la izquierda, y el usuario tiene que intervenir para indicar si se quiere continuar recto o tomar la salida. También se ocasiona la pérdida del carril ante oclusiones parciales de las marcas viales.
- Respecto a las condiciones medioambientales, el sistema funcionó correctamente durante el mediodía, con el Sol en alto, el amanecer y el atardecer, con el Sol en el horizonte, y de noche. Si embargo, sí se menciona como problema a resolver los reflejos del salpicadero que se producen en la luna delantera, y que por tanto son captados por la cámara.
- Otro problema fue la capacidad limitada para autodetectar el fallo del sistema. En las zonas de obras, sin marcas viales, el sistema pasaba automáticamente el control al conductor, avisándole con una señal sonora. No obstante, en algunos casos en que se perdió el rastro de la línea, por oclusiones debidas a otros vehículos, o por saturación de la imagen al salir de un túnel, fue el conductor el que tomó la iniciativa para evitar la salida de carril.

2.6.4. UTA II (DaimlerChrysler)

La DaimlerChrysler ha desarrollado el sistema UTA II (*“Urban Traffic Assistant”* en inglés), para la asistencia a la conducción en entornos urbanos (Vlacic y otros, 2001). Posee módulos de detección de obstáculos, peatones, carriles, señales de tráfico, semáforos, marcas viales, pasos de cebra, además de los módulos de clasificación de vehículos, control del vehículo e interfaz con el conductor.

Detección de carreteras

Las carreteras se construyen bajo el requisito de mantener cambios de curvatura suaves. Para ello, generalmente se sigue el modelo de clotoide que se muestra en la siguiente ecuación:

$$c(L) = c_0 + c_1 \cdot L \quad (2.4)$$

donde $c(L)$ describe la curvatura a lo largo de la longitud L de la curva, c_0 es la curvatura inicial, y c_1 la variación de curvatura, denominado parámetro de la clotoide, y que se considera constante en cada tramo. La curvatura se define como $c = \frac{1}{R}$, siendo R el radio de curvatura.

Assumiendo el modelo de óptica *pin-hole*, y conociendo la distancia focal f , el ángulo de inclinación α , y la altura H de la cámara, la relación entre un punto de la curva y su proyección sobre la imagen se puede describir mediante las siguientes ecuaciones:

$$\begin{cases} x_b = \frac{f}{L} \cdot \left(a \cdot w - x_{\text{off}} - \Delta\psi \cdot L + \frac{1}{2} \cdot c_0 \cdot L^2 + \frac{1}{6} \cdot c_1 \cdot L^3 \right) \\ L = \frac{H}{\alpha + \frac{y}{f}} \end{cases} \quad (2.5)$$

donde x_{off} , y $\Delta\psi$ definen respectivamente la posición lateral y la orientación del vehículo respecto del carril, w es el ancho del carril, y $a = \pm\frac{1}{2}$ se utiliza para referirse al borde izquierdo o derecho. Las funciones trigonométricas se han aproximado por su argumento ($\sin x \approx x$, $\tan x \approx x$).

Además, la conducción a altas velocidades exige tener en cuenta restricciones en la dinámica y cinemática del vehículo, que se expresan en forma de ecuaciones diferenciales. Para resolver el sistema de ecuaciones es necesario medir variables como la velocidad lateral causada por posibles deslizamientos, y la variación en la orientación del vehículo, lo que se consigue mediante sensores inerciales.

La integración de los modelos anteriores se realiza mediante el filtro de Kalman, aplicado a un modelo de variables de estado. Este filtro permite definir para la siguiente imagen un intervalo de búsqueda, delimitado por $\pm 3\sigma$, que en las distribuciones normales engloba al 99 % de los datos. El procesamiento de la imagen consiste en explorar una serie de líneas horizontales en la imagen, en busca de puntos de alto gradiente en la intensidad. Estas líneas de búsqueda tienen una anchura y posición extraídas de la estimación del filtro de Kalman.

Todo lo anterior se ha descrito asumiendo el modelo de carretera plana y es válido para un sistema de visión monocular. En ocasiones pueden detectarse como líneas viales reflejos sobre los guarda-raíles, si se encuentran dentro de la zona de exploración. Esto provoca el fallo del sistema, ya que afecta a la estimación de todas las variables.

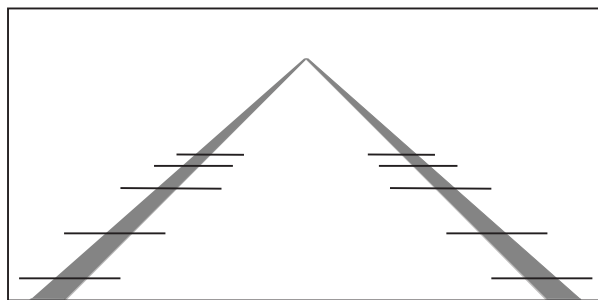


Figura 2.12:

Estos problemas pueden evitarse utilizando información estéreo, que permite, por una parte, descartar objetos que se salgan del plano de la carretera, y por otra, obtener la curvatura vertical de la carretera (Dickmanns y Mysliwetz, 1992). Ésta se supone independiente de la curvatura horizontal, y se aproxima también por una clotoide –ecuación (2.6)–, pero en este caso $c_{v,0}$ se considera constante, al igual que $c_{v,1}$.

$$c_v(L) = c_{v,0} + c_{v,1} \cdot L \quad (2.6)$$

Éste método de detección, como ya se comentó anteriormente, es altamente sensible a ruidos y desviaciones en los parámetros de calibración. Sirva como ejemplo el que de noche, para garantizar un funcionamiento óptimo, la región de carretera explorada no llega a tener 10m de longitud. Esto ha motivado la inclusión de otros dos sensores (un radar, y un navegador GPS con un mapa de carreteras), cuya información se integra con la de sistema de visión. De esta forma se mejora también el funcionamiento ante condiciones meteorológicas adversas en las que la visibilidad es muy limitada.

2.7. Sistemas comerciales

Comentamos aquí brevemente algunos sistemas comerciales de asistencia a la conducción que incluyen alerta por salida de carril utilizando visión por computador.

2.7.1. AutoVue, de Iteris

El primer sistema comercial de alerta por salida de carril en Europa es el denominado AutoVue (Iteris Inc., 2000), desarrollado por Iteris Inc. Se encuentra disponible en los camiones Mercedes Actros y MAN desde el año 2000. Se trata de una versión comercial del sistema anterior (apartado 2.6.4).

Iteris publica en su página web las siguientes características del sistema:

- No emite alertas por salidas de carril intencionadas (con los intermitentes activados).

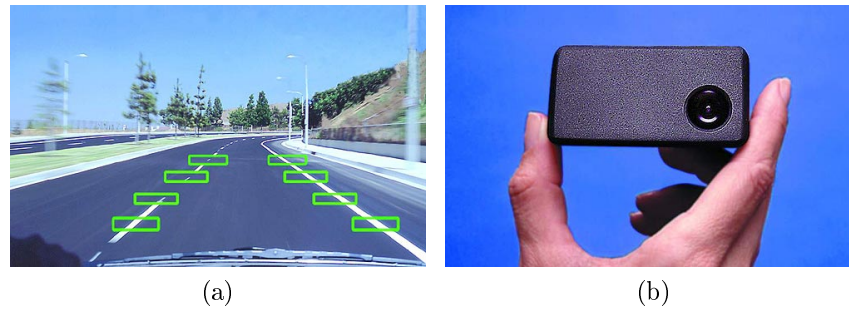


Figura 2.13: Sistema AutoVue de Iteris Inc.

- No necesita mantenimiento con limpiaparabrisas funcionales y luna delantera limpia.
- Se desactiva cuando la velocidad es inferior a 60km/h (diseñado para autovías/autopistas).
- Diseñado para una vida útil de 10-12 años.
- OEM disponible para Freightliner, Mercedes, y MAN.
-
- Muchos conductores de camión afirman que es efectivo con lluvia, agua-nieve, niebla, y en general en condiciones de visibilidad reducida.
- Las falsas alarmas son prácticamente inexistentes.
- Puede avisar de fatiga antes de que se convierta en un problema.

2.7.2. LDW de TRW

El sistema de Alerta por Salida de Carril de TRW ([TRW Automotive, ????](#)) utiliza una cámara frontal que monitoriza continuamente el carril por el que se circula. Si el conductor comienza a abandonar el carril de forma involuntaria, el sistema emite alertas visuales, sonoras, o por vibración del volante. Cuando el sistema se integra con el “Electrically Powered Steering”, también de TRW, éste es capaz de ejercer una ligera presión sobre el volante que ayuda al conductor a mantener el vehículo dentro del carril.

Según publicitan en su página web, el sistema funciona en condiciones diurnas, nocturnas, con sombras, a media luz, lluvia y niebla.

2.7.3. LDW de Mobileye

La empresa Mobileye ([Mobileye, ????](#)) ha desarrollado un sistema de asistencia a la conducción que incluye alerta por salida de carril.

En su página web, afirman detecta una gran variedad de marcas viales, incluyendo líneas continuas, líneas discontinuas, reflectores, líneas dobles , etc., bajo varias condiciones ambientales y tipos de carretera (hormigón y asfalto). El sistema mide la distancia de cada rueda a las líneas viales, y proporciona información del ancho y de la curvatura de la carretera.

La aplicación a la Alerta por Salida de carril se basa en el cálculo de la velocidad lateral del vehículo respecto a las líneas viales. El sistema puede ajustar los umbrales de alerta para adaptarlos al tipo de carretera, de forma que el aviso sea previo a la salida de carril, en el instante en que se pisa la línea, o permitir una cierta holgura para evitar falsas alarmas al recortar curvas en carreteras secundarias. También puede prevenir de somnolencia al conductor, cuando detecta patrones de conducción dentro del carril asociados a dicho estado.

La empresa publicó en agosto de 2007 que está suministrando tecnología a General Motors, para instalar en 2008 este sistema en los modelos Cadillac STS y Buick Lucerne Sedan. También va a suministrar, junto a Siemens VDO, un sistema de alerta por salida de carril para el nuevo BMW Serie 5.

2.7.4. Lane Departure Warning (LDW) / Lane Keeping System (LKS) de Continental AG

Continental AG, con su adquisición de la empresa Siemens VDO, adquirió también los sistemas de asistencia que ésta última estaba desarrollando. Dentro de estos sistemas se encuentra el “*Lane Departure Warning / Lane Keeping System (LDW / LKS)*”, que utiliza una cámara CMOS de alto rango dinámico, instalada cerca del espejo retrovisor interior, para detectar cuándo el vehículo se acerca a menos de una distancia prefijada de las líneas viales. El sistema se activa cuando la velocidad supera los 60km/h, y avisa de salidas de carril cuando no se activan los intermitentes, mediante un display en el panel de instrumentos, provocando vibraciones en el volante y ligeras correcciones en el volante para mantener el vehículo en el carril ([Continental Automotive, 2009](#)).

2.7.5. HELLA - DAS

La empresa HELLA ([HELLA, 2007](#)) también está actualmente desarrollando un sistema de aviso de salida de carril. .

El sistema emplea una cámara CMOS, integrada en el soporte del espejo retrovisor interior, que detecta la posición del vehículo dentro del carril. De nuevo, las alarmas pueden ser visuales, acústicas o hápticas (ligera fuerza correctiva sobre el volante), y se suprimen al activar los intermitentes.

2.7.6. AssistWare - SafeTRACK

La empresa ASSISTWARE TECHNOLOGY, recientemente adquirida por COGNEX ha desarrollado *SafeTRACK* ([Assistware Technology, 2006](#)), un sistema que genera alertas por falta de atención del conductor o por salidas de carril no intencionadas. Está pensado para ser utilizado en flotas de camiones, ya que incluye funciones de “caja negra”.

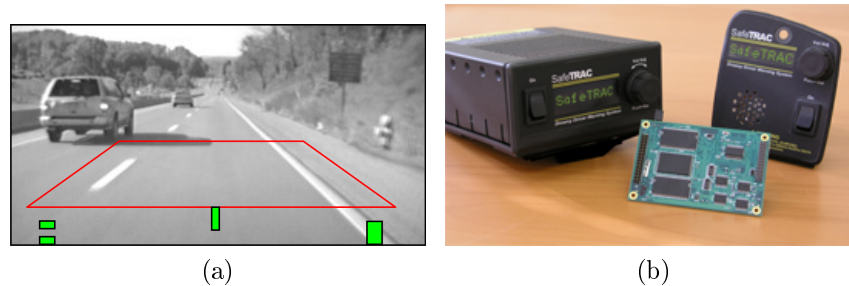


Figura 2.14: Sistema SafeTrack, de AssistWare

2.7.7. Toyota Lexus LS 460

Toyota ha incluido en su Lexus LS 460 ([Toyota, 2007](#)) un sistema de alerta por salida de carril que utiliza un sistema estéreo instalado detrás del espejo interior para monitorizar las líneas viales blancas (sujeto a las condiciones ambientales, clima y condiciones de la carretera, según comentan en la página web del producto).

El sistema, denominado “Lane Keeping Assist” ofrece dos funciones:

- “*Lane Departure Warning*” (Alerta por Salida de Carril): El sistema lanza una alerta audio-visual y aplica una ligera fuerza correctiva para mantener el vehículo en el carril.
- “*Lane Keep*” (Seguimiento de Carril): Proporciona una fuerza adicional en el volante para ayudar al conductor a mantener el vehículo en el carril.

2.7.8. Irisbus - CIVIS

Civis ([IrisbusIVECO, 2009](#)) es un sistema ligero de transporte urbano desarrollado conjuntamente por Irisbus y Siemens Transportation Systems, que pretende ofrecer un servicio equivalente al de los tranvías modernos. Integra un sistema de guiado óptico además de tracción eléctrica, pudiendo circular tanto por carril reservado como por la vía pública común. El sistema requiere un carril dedicado, pintado con dos líneas discontinuas y paralelas a escasa distancia, que marcan la trayectoria a seguir por el autocar. El

sistema de visión se compone de una cámara situada detrás del parabrisas (ver figura 2.15).

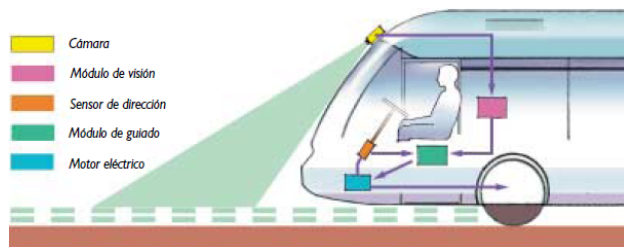


Figura 2.15: Irisbus IVECO Cavis

2.7.9. Honda Accord

Honda ha incluido en su gama Accord un paquete de Sistemas de Asistencia (figura 2.16), dentro de los cuales se encuentra el sistema LKAS (*Lane Keeping Assist System*).

Según la web del fabricante ([Honda, 2009](#)), el sistema LKAS funciona:

- en las carreteras donde las líneas de separación izquierda y derecha sean claras,
- en autopistas y carreteras similares,
- si el conductor mantiene un contacto correcto con el volante en todo momento (en caso contrario el sistema se desactiva automáticamente),
- si los intermitentes están apagados y la fuerza que emplea el conductor no es suficiente para indicar que podría querer cambiar de carril voluntariamente.



Figura 2.16: Algunos de los Sistemas ADAS del Honda Accord

2.8. Conclusiones

En sus orígenes, la investigación en sistemas de visión por computador dentro del sector del automóvil se orientó hacia los sistemas de conducción

automática (Turk y otros, 1988; Thorpe y otros, 1991; Dickmanns y Mysliwetz, 1992; Pomerleau, 1993; Broggi y otros, 1999). Fruto de ello son los sistemas descritos en el apartado 2.6, donde se han alcanzado tasas de conducción automática superiores al 90 %. Las pruebas intensivas realizadas en estos proyectos han puesto de manifiesto las dificultades encontradas, muchas de ellas aún sin resolver.

Sin embargo, desde hace poco más de una década, la investigación está derivando cada vez más hacia sistemas de asistencia a la conducción. En el apartado 2.5.2 se describen la diversidad de métodos desarrollados para reconocer el perfil de la carretera, ajustando modelos matemáticos cada vez más detallados.

Los sistemas de asistencia han alcanzado ya una robustez suficiente como para dar pie a su comercialización, generalmente destinada a vehículos de alta gama (Honda, 2009; Toyota, 2007), a flotas de camiones (Assistware Technology, 2006; Continental Automotive, 2009), e incluso a sistemas de conducción automática con un carril dedicado (IrisbusIVECO, 2009). En la actualidad, los sistemas comerciales de alerta por salida de carril funcionan exclusivamente en ciertos entornos (vías rápidas) y condiciones ambientales (buena visibilidad) cuando el vehículo supera aproximadamente los 60km/h. Son capaces de detectar los límites del carril por el que se circula, y percibir si el vehículo está en ese momento cruzando esos límites o no.

Resumen, los sistemas de asistencia existentes funcionan bien bajo ciertas condiciones, pero presentan aún las siguientes limitaciones:

- El reconocimiento del entorno, y por lo tanto, la capacidad de evaluación del riesgo son limitados:

Se reconocen la posición de las líneas viales, pero no se identifica su tipo, es decir, si se trata de una línea continua (no se debe cruzar nunca), o discontinua, que a su vez puede ser de varias clases. Es muy escaso el número de artículos publicados en relación a la clasificación de las distintas líneas viales (Campbell y Thomas, 1993; Risack y otros, 1998).

La mayoría de las técnicas desarrolladas recientemente siguen un enfoque descendente. Estos enfoques suelen ser más robustos que los ascendentes –facilitan el seguimiento (“tracking”) y son menos sensibles al ruido–, pero son incapaces de detectar objetos no contemplados en el modelo, como por ejemplo, el desdoblamiento de un carril en dos, o la aparición y desaparición de carriles adyacentes al propio. No ocurre así con los enfoques ascendentes, que podrían detectar todas las líneas viales visibles, pero dificultan las tareas de seguimiento y requieren etapas de postprocesado para eliminar las falsas detecciones.

Un sistema capaz de clasificar las líneas viales, así como de detectar el número y tipo de carriles de la carretera, podría distinguir entre dos situaciones tan distintas como el cruce de una línea discontinua

en una autovía de tres carriles y el cruce de una línea continua en una carretera convencional de doble sentido. Por lo tanto, dispondría de una capacidad de evaluación del riesgo mucho más precisa que los sistemas actuales.

- Anticipación a las maniobras y evaluación del riesgo.

La capacidad de anticipar maniobras peligrosas, como por ejemplo los adelantamientos indebidos, las salidas de carril no intencionadas, o la circulación a una velocidad excesiva, requieren, por una parte, el uso de modelos matemáticos más precisos, es decir, de mayor orden, y por otra, la coordinación de diferentes habilidades perceptivas (detección de carriles, vehículos, peatones, señales, etc.).

En resumen, se aprecian una serie de campos aún poco explorados en la investigación, relacionados con el reconocimiento de las marcas viales, el ajuste de modelos más completos del entorno de la carretera, y la integración de la detección de carreteras con otros módulos de detección. Con esta tesis se pretende profundizar en estas líneas de investigación, que se consideran fundamentales para mejorar las capacidades perceptivas de los vehículos inteligentes.

Capítulo 3

Autocalibración del sistema de visión estéreo

3.1. Introducción

El proceso de calibración permite al sistema establecer una correspondencia entre la posición de los objetos en el mundo real y su proyección en la imagen. Es por lo tanto un proceso imprescindible para poder obtener la posición relativa al vehículo de todos aquellos elementos detectados por el sistema de visión, que en el caso del IVVI son los peatones, otros vehículos, las señales de tráfico y las líneas de la carretera.

3.1.1. Parámetros de calibración

La correspondencia entre los sistemas de referencia asociados al mundo y a la imagen viene determinada por dos tipos de parámetros:

- Los **parámetros intrínsecos** son los relacionados con el conjunto de la cámara más la óptica. Generalmente se habla de:
 - *Centro de la imagen*: Punto en el que el eje óptico cruza el elemento sensor.
 - *Distancia focal*: Se suele dar una medida en horizontal (f_x) y otra en vertical (f_y), ya que las lentes no suelen ser perfectamente circulares.
 - *Distorsión*: Las imperfecciones de la lente provocan que la proyección sobre el elemento sensor modifica la geometría de la imagen. La distorsión geométrica puede ser:
 - radial: tipo *barril* o *cojín* (figura 3.1).
 - tangencial.

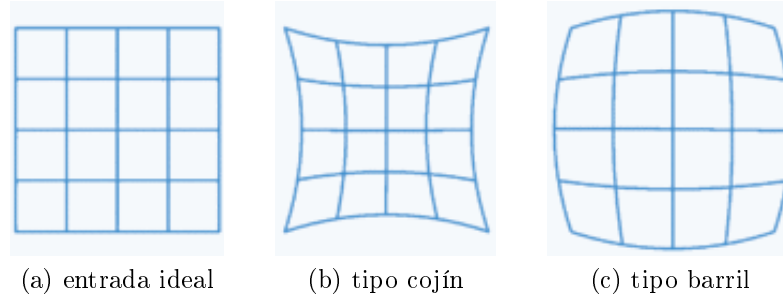


Figura 3.1: Tipos de distorsión radial

- Los **parámetros extrínsecos** se refieren a la posición y orientación del sistema de visión respecto del mundo (figura 3.2). Sean $OXYZ_{mundo}$ el Sistema de Referencia del Mundo y $OXYZ_{cámara}$ el Sistema de Referencia de la Cámara. Los parámetros extrínsecos *absolutos* vienen definidos por las matrices de rotación \mathbf{R}_{abs} y de traslación \mathbf{T}_{abs} , de forma que para traducir un punto expresado en el sistema de referencia del mundo (\mathbf{p}_{mundo}) al sistema de referencia de las cámaras ($\mathbf{p}_{cámara}$), se tiene:

$$\mathbf{p}_{cámara} = \mathbf{R}_{abs} \cdot \mathbf{p}_{mundo} + \mathbf{T}_{abs} \quad (3.1)$$

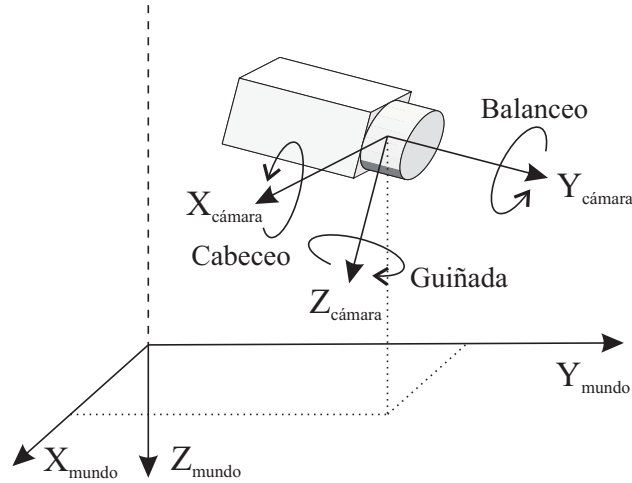


Figura 3.2: Parámetros extrínsecos

En el caso de un sistema estéreo o binocular, se necesita calibrar, además de éstos, los parámetros de transformación *relativa* (\mathbf{R}_{rel} y \mathbf{T}_{rel}) entre ambas cámaras (figura 3.3):

$$\mathbf{p}_{derecha} = \mathbf{R}_{rel} \cdot \mathbf{p}_{izquierda} + \mathbf{T}_{rel} \quad (3.2)$$

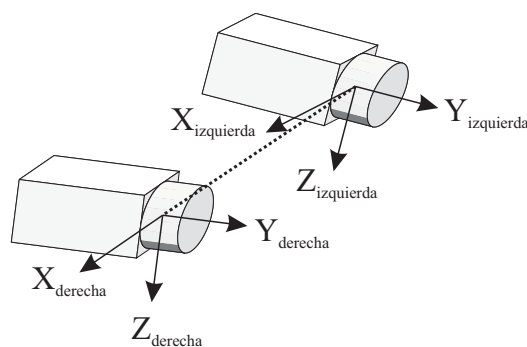


Figura 3.3: Sistema Estéreo

Existen varios métodos para la calibración de los parámetros intrínsecos de sistemas monoculares y binoculares. Una descripción de los más importantes puede encontrarse en (Faugeras, 1993). Dichos parámetros pueden ser obtenidos “fuera de línea”, y sin necesidad de tener el sistema de visión instalado en su posición definitiva.

Como es obvio, la calibración de los parámetros extrínsecos sólo puede realizarse con el sistema de visión ya montado en el vehículo en el que prestará servicio. Esto presenta varios inconvenientes:

- Por una parte, la gran variedad de vehículos, y por tanto, de posibles posiciones y orientaciones que puede tener el sistema de percepción, no hace factible la utilización de unos valores iniciales universales que se puedan ir ajustando mediante algún algoritmo recursivo.
- Por otra parte, es necesario un patrón de calibración visible desde el vehículo. La mayoría de los sistemas existentes requieren de patrones dibujados, ya sea en el suelo (Ernst y otros, 1999), en el capó del coche (Broggi y otros, 2001), o en un objeto móvil (Dang y Hoffmann, 2004).

3.1.2. Importancia de la calibración

En el caso de que el sistema de visión esté diseñado para ser integrado en un modelo concreto de vehículo, no hay nada que impida al fabricante realizar la calibración previa entrega al usuario. Sin embargo, si el sistema de visión no está destinado a un vehículo concreto, es entonces el proveedor del Sistema de Asistencia a la Conducción el encargado de proporcionar los mecanismos de calibración. En por lo tanto de vital importancia facilitar en lo posible al usuario la realización de la calibración extrínseca.

En este capítulo, nos referimos a la calibración manual como aquella que requiere de un patrón de calibración artificial, que deberá ser proporcionado por el fabricante. Por el contrario, entendemos por calibración automática aquella que no requiere de ningún patrón artificial. El principal problema que presenta la calibración manual es el tamaño del patrón, que deberá abarcar

un área similar a la región de la carretera analizada, y para que el sistema sea eficaz se requieren al menos 30-40m (McCall y Trivedi, 2006). Es fácil imaginar los inconvenientes que le puede suponer esto al usuario final.

Es por tanto necesario avanzar en el diseño de algoritmos que permitan realizar la calibración extrínseca de forma automática. A pesar de su importancia, este tema ha recibido hasta ahora una escasa atención, como lo manifiesta el reducido número de trabajos de investigación que lo tratan, y que se citarán en el apartado 3.1.3.

En este capítulo se propone un algoritmo de autocalibración desarrollado para aplicaciones de asistencia a la conducción, con el objetivo de facilitar la instalación del sistema de visión en vehículos comerciales. El algoritmo se ha diseñado bajo los siguientes requisitos:

- utilizar como patrón de calibración exclusivamente elementos de la infraestructura básica de la red de carreteras.
- permitir obtener el mayor número de parámetros extrínsecos posible.
- minimizar la intervención del usuario.

La intervención que este algoritmo exige al usuario se limita a la conducción del vehículo por un tramo de carretera de ciertas características (ha de ser recto, plano y con líneas viales marcadas) y activar mediante un pulsador el proceso de autocalibrado. A partir de entonces el sistema podrá comenzar a utilizarse, y sólo será preciso repetir el proceso si se decide reinstalar el sistema de visión en una posición diferente.

3.1.3. Trabajos Previos

En esta sección se describen aplicaciones de técnicas de calibración y autocalibración aplicadas a los sistemas de visión embarcados en vehículos.

Calibración de sistemas monoculares

La calibración de sistemas monoculares habitualmente requiere información explícita sobre el entorno, o bien, simplificar el problema reduciendo el número de parámetros que se quieren estimar. En (Bücher, 2000) el problema se reduce a estimar la altura e inclinación vertical de la cámara. Para estimar la inclinación vertical, es suficiente con localizar la línea de horizonte, pero para poder realizar medidas de distancia en la imagen, además, el usuario debe seleccionar dos puntos en la imagen e indicar la distancia real entre ellos.

De modo parecido, en (Southall y Taylor, 2001) se describe un método supervisado de calibración que requiere que el usuario introduzca la anchura del carril por el que circula el vehículo, e indique cuales de las líneas detectadas en la imagen delimitan dicho carril. Al igual que en (Bücher, 2000), solamente se considera la inclinación vertical de la cámara.

Otras técnicas utilizan varios fotogramas capturados en posiciones diferentes. En (Lee y Deng, 1991) se estima la orientación de una cámara en movimiento analizando secuencias de 2 ó 3 fotogramas. Se asume que la cámara no modifica su orientación durante el movimiento, y que tanto su altura como la distancia recorrida por el vehículo entre dos fotogramas consecutivos son conocidas.

Calibración de sistemas binoculares

Para la calibración de sistemas binoculares, habitualmente se utiliza un patrón colocado o impreso en el suelo. En (Ernst y otros, 1999) se calibra cada cámara por separado a partir una rejilla pintada en el pavimento. A partir de las calibraciones individuales obtiene la posición y orientación relativa de una cámara respecto de la otra, así como la posición y orientación del sistema estéreo respecto del mundo. Para ello se parte de un sistema de ecuaciones linealizado que se resuelve mediante mínimos cuadrados recursivos.

Otros trabajos incluyen etapas supervisadas. En (Broggi y otros, 2001) se utiliza también una rejilla pintada en el suelo. Este patrón se captura en una imagen sobre la cual un usuario debe seleccionar a mano los puntos de intersección de las líneas de la trama. Con esto se consigue un ajuste “grueso” de los parámetros. Después, en una etapa no supervisada, la estimación se refina mediante un proceso iterativo, bajo la hipótesis de que la imagen capturada proviene de una superficie plana y con textura. Esta técnica se utiliza también para corregir pequeñas desviaciones en los parámetros extrínsecos originadas por el movimiento y vibraciones del vehículo, utilizando en este caso un conjunto de marcas colocadas en el capó, y que son fácilmente detectables por el sistema de visión. Esto permite corregir las desviaciones en los parámetros respecto del coche, pero no respecto de la carretera.

En (Labayrade y Aubert, 2003), se utiliza el algoritmo “v-disparity” para estimar de forma secuencial los ángulos de cabeceo, balanceo y guiñada (en inglés *roll*, *pitch* y *yaw*), sin requerir un patrón de calibración artificial. Realiza experimentos con imágenes sintéticas, obteniendo precisiones superiores a los $0,5^\circ$ en la estimación de los ángulos. Estas precisiones son posibles gracias a un sistema estéreo cuya distancia entre cámaras es de 1,02 m. En las pruebas con secuencias reales, obtiene valores coherentes, pero al no disponer de valores reales de los parámetros, no puede ofrecer una medida de precisión.

En (Sappa y otros, 2006) se propone un método que extrae la posición 3D de varios puntos de la imagen pertenecientes a la carretera, a los que ajusta un plano mediante el algoritmo RANSAC. Se considera que pertenecen a la carretera aquellos grupos de puntos suficientemente numerosos y cuya coordenada vertical es la más alejada de las cámaras mirando hacia abajo. Obtiene así la posición relativa de las cámaras respecto del plano de la carretera, partiendo de un sistema estéreo del que se conocen los parámetros intrínsecos y la posición relativa entre cámaras. El algoritmo funciona a unos 350ms por fotograma en un Pentium-IV a 3,2GHz.

En (Nedevschi y otros, 2006) se estiman los parámetros de forma secuencial. Utiliza las marcas viales como patrón, por lo que la etapa de segmentación es más compleja. Las marcas viales se analizan en la imagen de la cámara derecha, mediante la exploración del gradiente de la imagen y la aplicación una serie de restricciones geométricas a los bordes detectados. En el primer fotograma se aplica un algoritmo de detección del punto de fuga, y en los siguientes se aplica un algoritmo de seguimiento. Una vez detectado el punto de fuga, se filtran las marcas viales mediante la construcción de un histograma polar centrado en el punto de fuga. Se excluyen los bordes cuya orientación difiere demasiado de los máximos locales del histograma polar. A continuación, se calculan mediante mínimos cuadrados ponderados, y de forma secuencial, primero la orientación del eje vertical, después la orientación del eje transversal, y la altura, y por último la orientación del eje longitudinal. El algoritmo completo se evalúa en una secuencia de 50 fotogramas, en la que se obtienen valores coherentes y próximos a los calibrados en estático. Entre las conclusiones se menciona que la estimación de la guiñada (“yaw” en inglés) depende mucho de la habilidad del conductor para mantener el vehículo paralelo a las líneas. La ejecución dura 4,1 ms en un Pentium-IV a 2,6 GHz.

3.1.4. Breve Descripción

En este capítulo se presenta un algoritmo de autocalibración basado en un nuevo enfoque, que permite estimar simultáneamente la altura y los ángulos de cabeceo y balanceo del sistema de visión estéreo, sin la necesidad de un patrón de calibración artificial. La figura 3.4 muestra un esquema general del algoritmo propuesto.

En una primera etapa se detectan las líneas viales que delimitan el carril por el que circula el vehículo mediante la aplicación de un filtrado espacial y la transformada de Hough. Esta transformada requiere que las líneas viales sean rectas, pero tiene como ventaja que al ser un método global es robusto frente a discontinuidades en las líneas, permitiendo que las líneas viales utilizadas como patrón sean de cualquier tipo (continuas, discontinuas de separación de carriles normales, discontinuas de separación de carriles de entrada/salida, etc.).

La segunda etapa consiste en la selección de un conjunto de puntos pertenecientes a las líneas detectadas. Este conjunto constituirá el patrón de calibración.

Por último, los parámetros extrínsecos se estiman maximizando una función de salud mediante un algoritmo genético (Goldberg, 1989). Cada individuo representa una solución (altura y orientación del sistema estéreo) cuya “salud” se evalúa con una función que involucra tres etapas:

1. Primero se realiza una transformación de perspectiva que invierte el proceso de proyección de la imagen sobre la CCD. Es decir, se “proyec-

ta”, sobre el plano de la carretera, el patrón de puntos extraído de la imagen capturada por la CCD.

2. Después, se miden el paralelismo de las líneas viales y la proximidad entre los dos patrones proyectados desde ambas cámaras.
3. Finalmente se genera un valor de “salud” a partir de las medidas anteriores, que representa la coherencia entre los patrones proyectados desde la cámara izquierda y derecha respectivamente.

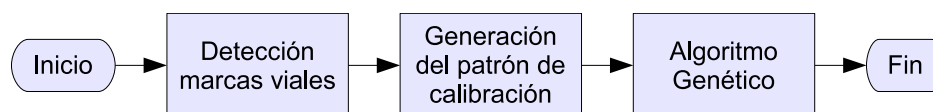


Figura 3.4: Diagrama de flujo

3.2. Descripción del Algoritmo

Las líneas viales de la carretera poseen unas características conocidas:

- Son paralelas.
- Existen tramos en que son rectas.
- En tramos rectos están alineadas con el vehículo.

El algoritmo de autocalibración propuesto parte de esta información del entorno. Su único prerequisite es la circulación por un tramo recto, plano, y con líneas viales marcadas, como el caso que se muestra en la figura 3.5-a. Este entorno es capturado por las dos cámaras del sistema estéreo dando lugar a las imágenes de la figura 3.5-b que serán procesadas para detectar las dos líneas viales que delimitan el carril, y que aparecen superpuestas en color rojo (cámara derecha) y verde (cámara izquierda).

Ya que la posición y orientación de las cámaras es desconocida, mediante un algoritmo genético se generan múltiples hipótesis que se combinarán entre sí mediante las operaciones de mutación y sobrecruzamiento. Para cada hipótesis (figura 3.5-c), se invierte la transformación de perspectiva de modo que las líneas detectadas en las imágenes capturadas se proyectan sobre el supuesto plano de la carretera (figura 3.5-d). En dicho plano, las líneas viales proyectadas desde ambas cámaras deberán coincidir y ser paralelas en, y sólo en, el caso de que los parámetros hipotéticos coincidan con los reales. El algoritmo genético evalúa cada hipótesis en base a los dos siguientes criterios:

- cuanto mayor sea el paralelismo entre las líneas, mejor es la solución.

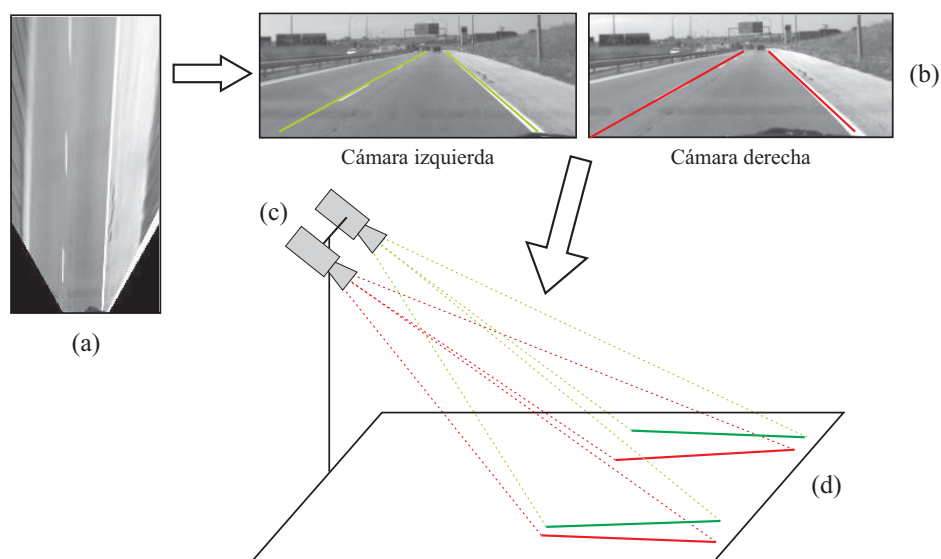


Figura 3.5: Descripción del Algoritmo

- cuanto menor sea la distancia entre las parejas de líneas proyectadas desde cada cámara, mejor es la solución.

A continuación se detalla el modelo matemático del sistema de visión estéreo, que permitirá realizar las transformaciones de perspectiva requeridas por el algoritmo.

3.2.1. Modelo del Sistema de Visión Estéreo

El proceso de calibración intrínseca, realizado con la *toolbox* de Matlab desarrollada en (Bouguet, 2006), permite que las imágenes capturadas puedan rectificarse para eliminar la distorsión y corregir las pequeñas diferencias entre ambas cámaras (apéndice B). De este modo se obtienen dos imágenes como si éstas proviniesen de dos cámaras idénticas y perfectamente alineadas. Gracias a la rectificación, puede considerarse el modelo del sistema estéreo como un sistema compuesto por dos cámaras paralelas y separadas una distancia fija.

A este modelo se le asocia un sistema de coordenadas cuyo origen está en el sensor de la cámara izquierda, y cuya orientación es la que se muestra en la figura 3.6). Se ha tomado como referencia la cámara izquierda ya que ésta será la utilizada posteriormente para la detección de carriles.

La proyección de cualquier punto de la carretera sobre cada CCD se puede calcular mediante matrices homogéneas, que permiten representar transformaciones complejas con una sola matriz. Para obtener la matriz que representa la transformación de perspectiva es necesario seguir los siguientes pasos:

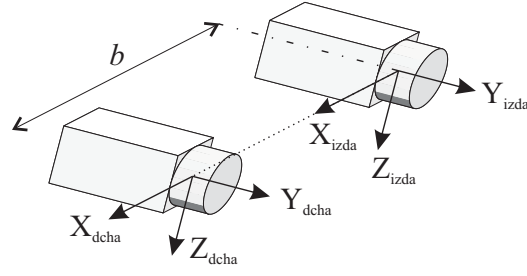


Figura 3.6: Modelo del sistema de visión estéreo rectificado

1. El sistema de coordenadas se eleva una altura h (figura 3.7a):

$$\mathbf{x}_1 = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & h \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{T_h} \cdot \mathbf{x}_0 \quad (3.3)$$

2. El sistema de coordenadas se rota un ángulo de cabeceo $-\alpha$ y un ángulo de balanceo β sobre los ejes OX y OZ , respectivamente (figuras 3.7b y 3.7c). Al ángulo de cabeceo se le invierte el signo para que el sentido positivo indique un giro en sentido horario, es decir, con las cámaras rotando para orientarse hacia al suelo, en vez de hacia el cielo. El ángulo de guiñada se calculará a posteriori, por lo que no se contempla por el momento. :

$$\mathbf{x}_2 = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha & 0 \\ 0 & \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{R_\alpha} \cdot \mathbf{x}_1 \quad (3.4)$$

$$\mathbf{x}_3 = \underbrace{\begin{bmatrix} \cos \beta & \sin \beta & 0 & 0 \\ -\sin \beta & \cos \beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{R_\beta} \cdot \mathbf{x}_2 \quad (3.5)$$

3. Sólo para el caso de la cámara derecha, el sistema de coordenadas se desplaza sobre el eje OX una distancia b (del inglés *baseline*), ya que el proceso de rectificación permite suponer que las cámaras están ali-

neadas (figura 3.7d):

$$\left\{ \begin{array}{l} \mathbf{x}_{c,dcha} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & -b \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{T_{b,dcha}} \cdot \mathbf{x}_3 \\ \mathbf{x}_{c,izda} = \underbrace{\mathbf{I}}_{T_{b,izda}} \cdot \mathbf{x}_3 \end{array} \right. \quad (3.6)$$

4. Se proyecta la imagen a lo largo del eje OZ con una distancia focal f , según el modelo *pin-hole*. El modelo pin-hole tiene la ventaja de poder expresarse como transformación lineal en coordenadas homogéneas, y eliminar la indeterminación de los puntos en el infinito, por lo que es ampliamente utilizado en geometría descriptiva.

$$\underbrace{\begin{bmatrix} u_p \\ v_p \\ s_p \end{bmatrix}}_{\mathbf{u}_p} = \underbrace{\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{T_f} \cdot \underbrace{\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}}_{\mathbf{x}_c} \quad \text{donde} \quad \begin{cases} x_p = \frac{u_p}{s_p} \\ y_p = \frac{v_p}{s_p} \end{cases} \quad (3.7)$$

5. Las coordenadas se traducen de milímetros a píxeles, y el origen de coordenadas se traslada desde el centro óptico de la CCD a la esquina superior izquierda de la imagen (figura 3.7e):

$$\mathbf{u}_i = \underbrace{\begin{bmatrix} k_x & 0 & c_x \\ 0 & k_y & c_y \\ 0 & 0 & 1 \end{bmatrix}}_{T_k} \cdot \mathbf{u}_p \quad (3.8)$$

donde k_x y k_y se refieren a la resolución espacial, expresada en píxeles/mm, y c_x y c_y son las coordenadas del centro óptico de la CCD, expresadas en píxeles.

Multiplicando estas matrices en orden inverso al presentado, se obtiene la matriz de transformación global:

$$\mathbf{u}_i = \begin{bmatrix} x_i \\ y_i \\ s_i \end{bmatrix} = (T_k \cdot T_f \cdot T_b \cdot R_\beta \cdot R_\alpha \cdot T_h) \cdot \mathbf{x}_m = T_{\text{global}} \cdot \begin{bmatrix} x_m \\ y_m \\ z_m \\ 1 \end{bmatrix} \quad (3.9)$$

Esta matriz T_{global} es de dimensión 3×4 . Si aplicamos la hipótesis de carretera plana, tendremos que $y_i = 0$, de modo que se puede prescindir

de la segunda columna de esta matriz. Llegamos así a una matriz T'_{global} (ecuación 3.10) que relaciona los puntos del plano de la carretera (x_m, z_m) con los del plano de la cámara (x_i, y_i) (ecuación 3.11). Cada cámara tendrá su matriz correspondiente, las cuales difieren solamente en el paso (3), y en los valores de sus centros (c_x, c_y) .

$$T'_{\text{global}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot T_{\text{global}} \quad (3.10)$$

$$\begin{bmatrix} u_i \\ v_i \\ s_i \end{bmatrix} = T'_{\text{global}} \cdot \begin{bmatrix} x_m \\ z_m \\ 1 \end{bmatrix} \quad \text{donde} \quad \begin{cases} x_i = \frac{u_i}{s_i} \\ y_i = \frac{v_i}{s_i} \end{cases} \quad (3.11)$$

3.2.2. Obtención del patrón de calibración

El patrón de calibración se obtiene en tres etapas:

Primera Etapa: Detección de marcas viales

Esta etapa comienza con el cálculo del gradiente horizontal de la imagen. Para ello, se realiza un filtrado con la máscara correspondiente a la ecuación 3.12. Este filtro, aplicado a la imagen de la figura 3.8a, da como resultado la figura 3.8b.

$$\frac{\Delta f(x)}{\Delta x} = \frac{f(x+1) - f(x-1)}{2} \iff \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \quad (3.12)$$

Si en estas imágenes extraemos el perfil de intensidades a lo largo de una horizontal, como por ejemplo, la línea blanca de la figura 3.9, se aprecia cómo las marcas viales aparecen como bandas blancas sobre un fondo oscuro. En el perfil del gradiente horizontal de intensidades (la misma figura, abajo), los bordes interno y externo de cada línea dan lugar a un máximo local seguido de un mínimo local, ambos muy próximos entre sí y situados en la misma fila de la imagen. Este patrón (un máximo local seguido de un mínimo local a una corta distancia) es el que indica la presencia de una marca vial.

Este mecanismo de detección de marcas es muy similar al que se utiliza en la detección y seguimiento de carriles (capítulo 4), con dos salvedades. La primera, que aquí se aplica a la imagen en perspectiva de la cámara (figura 3.5b) en lugar de a la imagen en perspectiva inversa (figura 3.5a). La segunda, que en la autocalibración, de los dos bordes de la marca vial se considera el del lado izquierdo, correspondiente al máximo local en el perfil de gradiente y que se muestra en la figura 3.9 con dos líneas discontinuas, a diferencia de la detección y seguimiento de carriles en la que se considera el píxel central de la marca vial, equidistante a ambos bordes de la línea (como

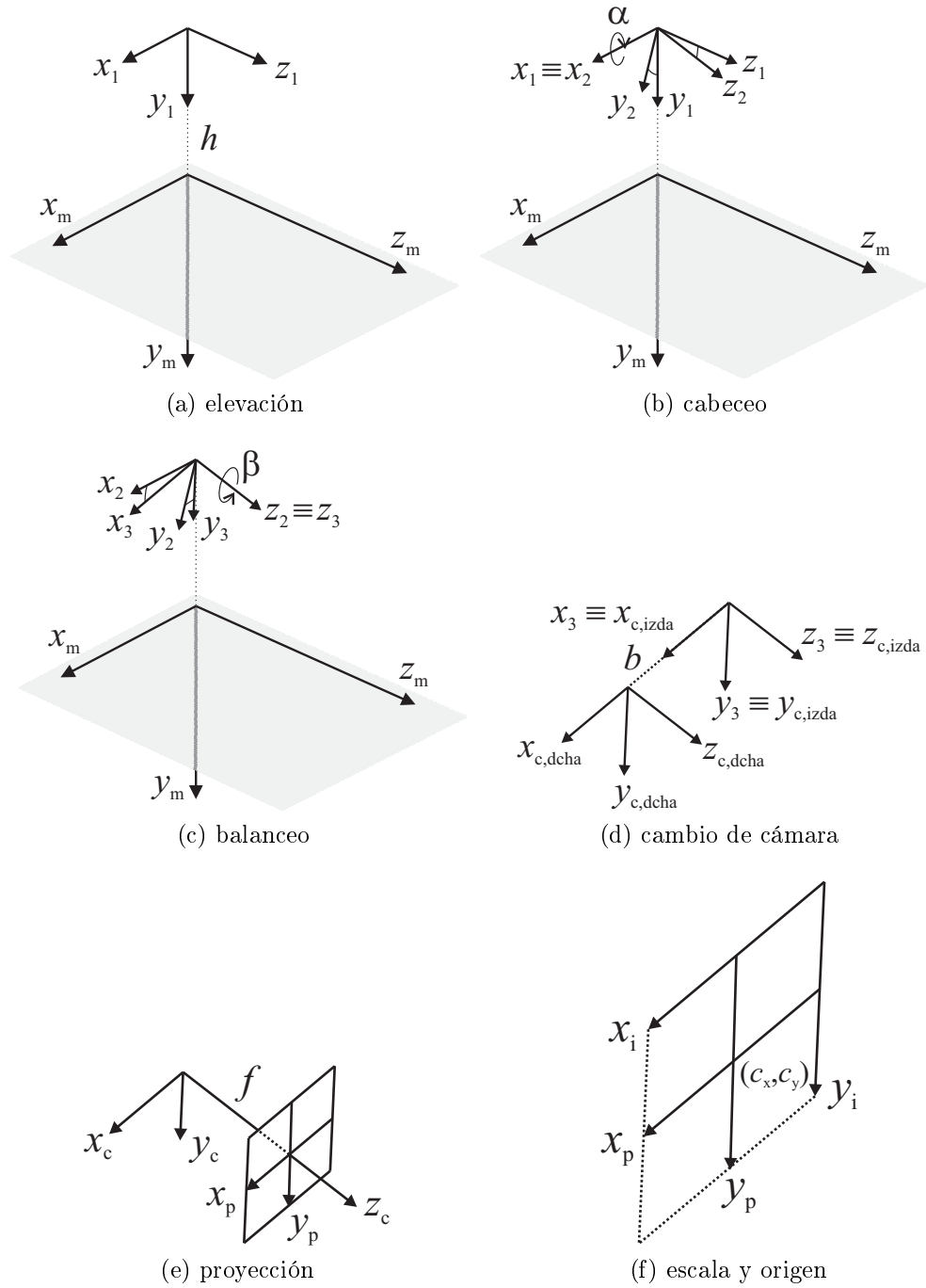
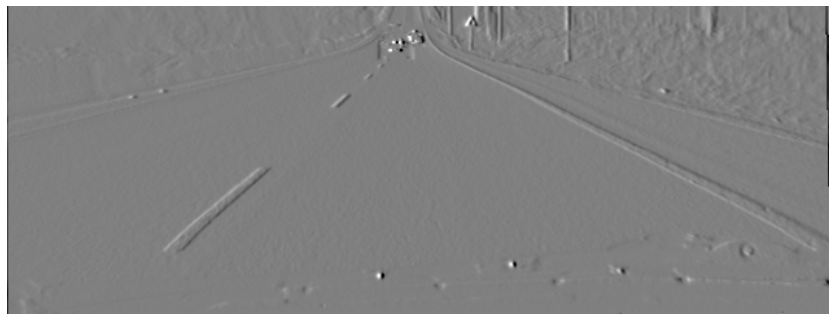


Figura 3.7: Cambio de Sistema de Referencia



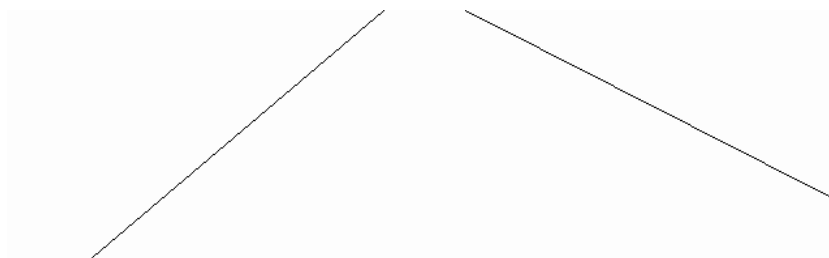
(a) Carretera



(b) Imagen filtrada



(c) Detección de marcas viales



(d) Detección de líneas viales con la Transformada de Hough

Figura 3.8: Detección de líneas viales para la generación del patrón de calibración

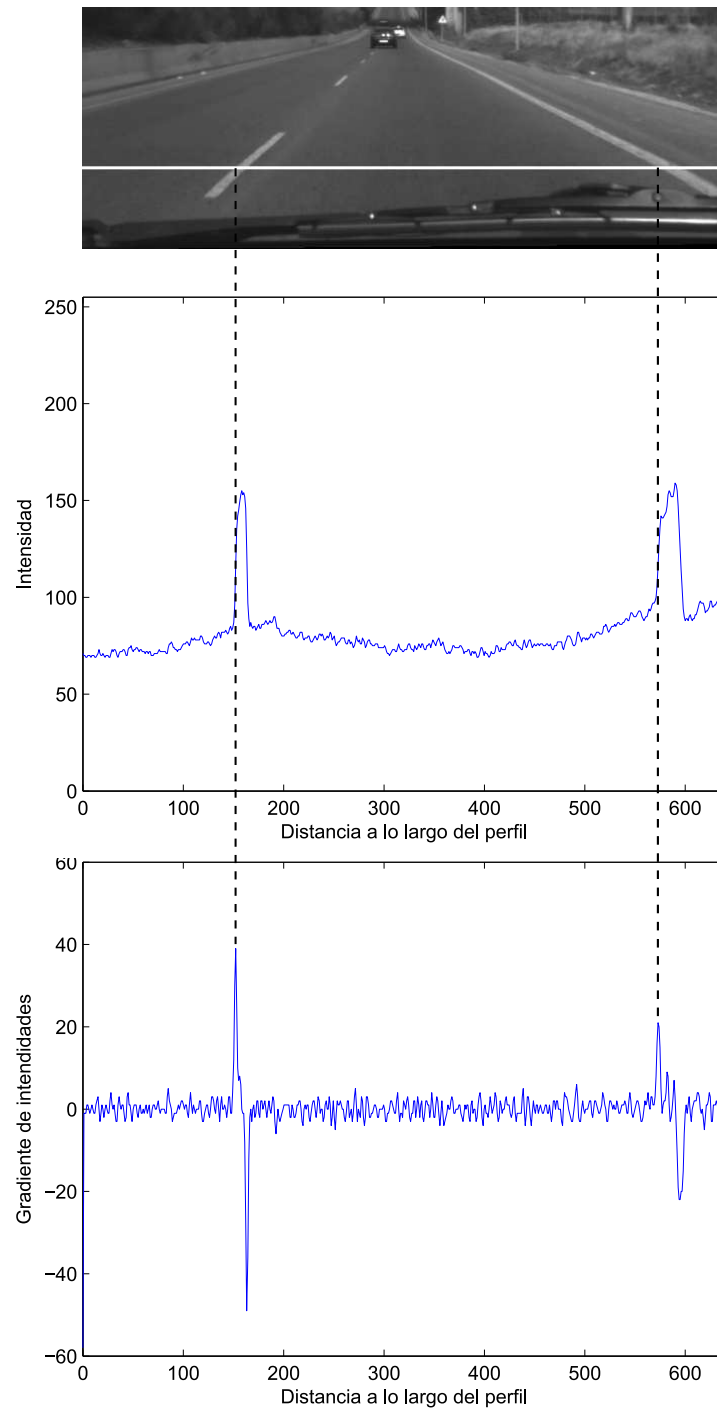


Figura 3.9: Detección de marcas viales basada en el gradiente

se explicará en el apartado 4.3). Esto es así para evitar que el patrón sea dependiente de la perspectiva. La figura 3.10 ilustra este fenómeno que se produjo en las primeras versiones del algoritmo. Cuando se considera el píxel central de las líneas viales, el ángulo de balanceo (exagerado en la figura para apreciar bien el efecto) hace que el píxel central marcado sobre una horizontal, y el eje real de la línea vial estén separados una distancia d_i o d_d .

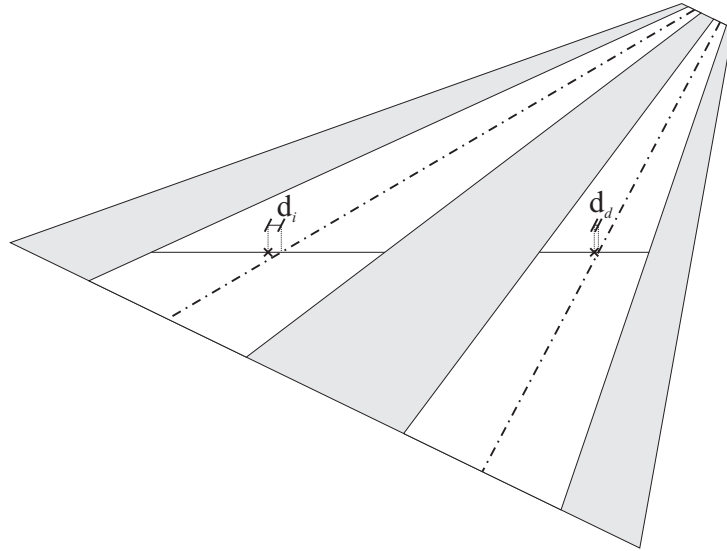


Figura 3.10: Efecto de la perspectiva sobre la detección de marcas viales

Segunda Etapa: Transformada de Hough

En la segunda etapa se extraen las líneas izquierda y derecha del carril aplicando la Transformada de Hough (Hough, 1959) al conjunto de píxeles marcados en la etapa previa. En el acumulador de Hough, se delimitan dos regiones de búsqueda, una para la línea izquierda y otra para la derecha. Dentro de cada región se selecciona la línea más votada. El resultado, para una cámara, se muestra en la figura 3.8d.

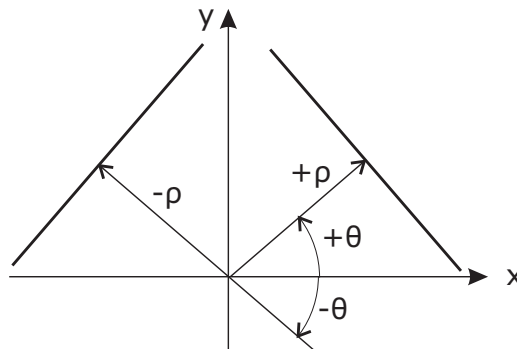


Figura 3.11: Parametrización de Hough: $\rho = x \cos \vartheta + y \sin \vartheta$

Al proyectar estas líneas desde la cámara sobre el plano de la carretera (figura 3.5c), se obtiene una pareja de líneas proyectadas desde la cámara izquierda y otra desde la derecha. Utilizando la parametrización de Hough (figura 3.11), cada línea vendrá definida por la pareja de parámetros (ρ, θ) . Para referirnos a cada una de ellas, denominaremos con los subíndices ci y cd a las líneas procedentes de la cámara izquierda y derecha, respectivamente, mientras que distinguiremos las líneas izquierda y derecha detectadas en la misma cámara mediante los subíndices li y ld :

	línea izquierda	línea derecha
cámara izquierda	$(\rho_{ci,li}, \vartheta_{ci,li})$	$(\rho_{ci,li}, \vartheta_{ci,ld})$
cámara derecha	$(\rho_{cd,li}, \vartheta_{cd,li})$	$(\rho_{cd,li}, \vartheta_{cd,ld})$

A continuación debe definirse un mecanismo para evaluar su proximidad y su paralelismo, de acuerdo a los criterios mencionados al principio del apartado 3.2. La manera más evidente es comparar directamente sus parámetros. Si se calculan las diferencias entre los parámetros de las líneas izquierda y derecha (ecuación 3.13), es posible obtener una medida de la discrepancia entre ambas mediante la ecuación 3.14, donde k_ρ y k_θ son constantes que permiten variar la importancia relativa de cada término.

$$\begin{aligned} d_{\rho,izda} &= |\rho_{ci,li} - \rho_{cd,li}|; & d_{\rho,dcha} &= |\rho_{ci,ld} - \rho_{cd,ld}| \\ d_{\theta,izda} &= |\vartheta_{ci,li} - \vartheta_{cd,li}|; & d_{\theta,dcha} &= |\vartheta_{ci,ld} - \vartheta_{cd,ld}| \end{aligned} \quad (3.13)$$

$$\text{error} = k_\rho (d_{\rho,izda} + d_{\rho,dcha}) + k_\theta (d_{\theta,izda} + d_{\theta,dcha}) \quad (3.14)$$

Sin embargo, este mecanismo de comparación nunca obtuvo una adecuada convergencia del algoritmo genético, ni con la parametrización ρ - θ , ni con la parametrización clásica $y = mx + b$. Por este motivo se añadió una tercera etapa, que extrae de estas líneas un conjunto de puntos.

Tercera Etapa: Patrón de Calibración

Antes de ser proyectadas sobre el plano de la carretera, las líneas detectadas se evalúan a diferentes alturas en la imagen, espaciadas 10 píxeles, comenzando por la fila inferior y subiendo hasta alcanzar el 90 % de la altura del punto de fuga (figura 3.12).

Se obtiene así un patrón de puntos para cada cámara. Estos patrones de puntos son proyectados sobre el plano de la carretera, y comparados entre sí mediante la función de salud del algoritmo genético.

3.2.3. Función de Salud del Algoritmo Genético

Cada individuo de la población del algoritmo genético representa una posición y orientación del sistema estéreo. La salud asociada a cada individuo deberá evaluar que:

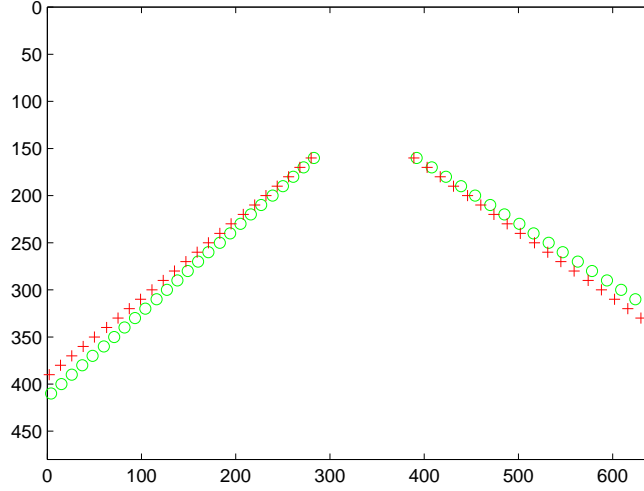


Figura 3.12: Patrón de calibración; (+) patrón desde la cámara derecha; (o) patrón desde la cámara izquierda

- las proyecciones de los patrones, desde ambas cámaras, sobre el plano de la carretera deben formar dos conjuntos de puntos coincidentes,
- estos patrones forman sobre la carretera un conjunto de líneas paralelas.

Estas dos condiciones serán evaluadas por una función de error compuesta por dos términos, E_1 y E_2 :

Primer término

Sean $\vec{x}_{i, izquierda}$ y $\vec{x}_{i, derecha}$ los conjuntos de puntos correspondientes a las imágenes izquierda y derecha respectivamente, obtenidos tal como se ha explicado en el apartado 3.2.2. El primer término de la función de error evalúa que cada punto proyectado sobre el plano de la carretera desde una cámara, coincide con la proyección del mismo punto desde la otra cámara. Se toma como error la suma de los cuadrados de las distancias euclídeas entre las proyecciones de cada punto obtenida desde ambas cámaras:

$$E_1 = \sum_i \|\vec{x}_{i, izquierda} - \vec{x}_{i, derecha}\|^2 \quad (3.15)$$

Segundo término

El segundo término evalúa el paralelismo entre las líneas viales. Sea ϑ el ángulo que forma la línea vial con el eje horizontal. Los subíndices 1 y 2 representan respectivamente las líneas izquierda y derecha del carril, mientras que los subíndices *dcha* e *izda* hacen referencia a la cámara.

El error en paralelismo se calcula como el mínimo entre la diferencia angular de las líneas de la cámara derecha y la diferencia angular de las de la cámara izquierda, según la ecuación 3.16:

$$E_2 = \min(|\vartheta_{izda,1} - \vartheta_{izda,2}|, |\vartheta_{dcha,1} - \vartheta_{dcha,2}|) \quad (3.16)$$

El motivo de tomar el valor mínimo, en lugar de la suma (ecuación 3.17) o del máximo (ecuación 3.18), como parecería más lógico, responde a que estas ecuaciones producen valores de error elevados incluso en las inmediaciones de la solución correcta. Aunque la función *mínimo* solo fomenta la búsqueda del paralelismo en las líneas proyectadas desde una de las cámaras, utilizada en combinación con el término anterior (ecuación 3.15) consigue que los conjuntos de puntos de ambas cámaras tiendan a la vez a alinearse y aproximarse.

$$E_2' = |\vartheta_{izquierda,1} - \vartheta_{izquierda,2}| + |\vartheta_{derecha,1} - \vartheta_{derecha,2}| \quad (3.17)$$

$$E_2'' = \max(|\vartheta_{izquierda,1} - \vartheta_{izquierda,2}|, |\vartheta_{derecha,1} - \vartheta_{derecha,2}|) \quad (3.18)$$

Función completa

El siguiente paso consiste en definir una función de salud a partir de los términos de error E_1 y E_2 .

Para equilibrar el peso de ambos términos en la función global, y para incrementar la resolución de la función de salud cuando ésta se aproxime a la solución real, se ha utilizado la función inversa para normalizar y transformar E_1 y E_2 en S_1 y S_2 , según las ecuaciones 3.19, donde K_1 y K_2 son dos constantes que controlan el perfil de la curva y acotan el valor de cada término al intervalo $(0, 1)$.

$$\begin{cases} S_1 = \frac{K_1}{E_1 + K_1} \\ S_2 = \frac{K_2}{E_2 + K_2} \end{cases} \quad (3.19)$$

Una vez definidos los términos S_1 y S_2 , sólo queda expresar la función de salud como combinación de ambos. La mejor solución encontrada consiste en restar el valor mínimo de ambos términos al valor máximo de la función (en este caso es igual a 1), según la ecuación 3.20. Por razones similares al caso anterior, se descartó la función suma ($Salud = S_1 + S_2$) por que tiende a maximizar uno de los términos en detrimento del otro, y sólo da buenos resultados con imágenes ideales, es decir, sintéticas y con ausencia total de ruido.

$$Salud = 1 - \min(S_1, S_2) \quad (3.20)$$

3.3. Resultados

3.3.1. Con Imágenes Sintéticas

Para evaluar la eficacia del algoritmo se han utilizado imágenes sintéticas, ya que éstas permiten comparar las estimaciones con unos parámetros reales que se conocen con total precisión. La figura 3.13 muestra el modelo de carretera utilizado, y las proyecciones sobre las dos cámaras.

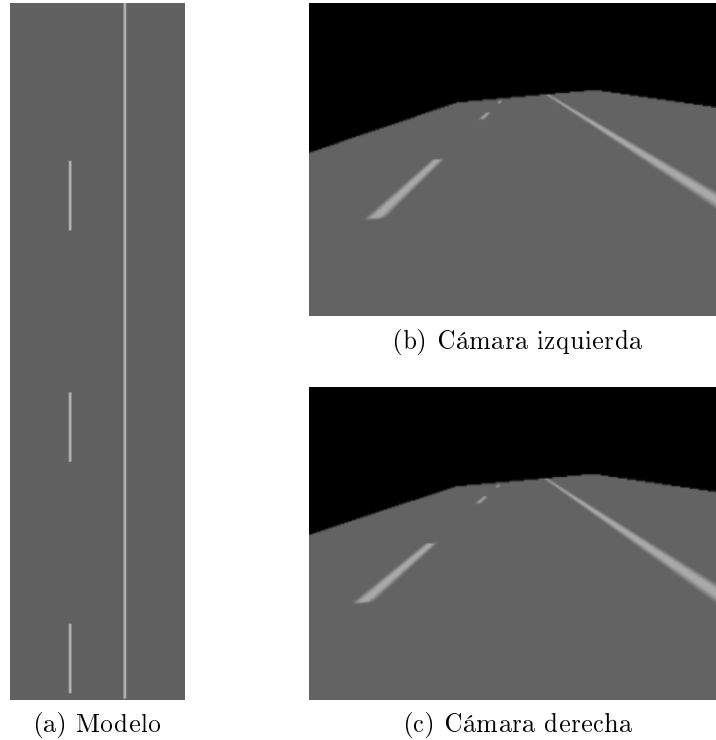


Figura 3.13: Modelo de carretera proyectado sobre las cámaras del sistema de visión estéreo

Las transformaciones de perspectiva se han realizado con los parámetros intrínsecos mostrados en la tabla 3.1, que coinciden con los de las cámaras embarcadas en el vehículo IVVI.

La tabla 3.2 muestra los parámetros del algoritmo genético (AG), mientras que la tabla 3.3 presenta los parámetros estimados comparados con los reales, así como la varianza de la estimación calculada a partir de diez ejecuciones del algoritmo. Esta varianza es del orden de 10^{-2} en altura, 10^{-8} en el ángulo de cabeceo, y 10^{-10} en el de guiñada, lo cual demuestra la convergencia del algoritmo.

La figura 3.14 muestra un ejemplo gráfico de resultados. La gráfica de la izquierda representa la evolución del error, para comprobar la convergencia. En todas las ejecuciones del algoritmo, éste convergió antes de 20 generaciones. En la misma figura, la gráfica de la derecha muestra la correspondencia entre el patrón de puntos proyectado desde la cámara derecha (con “+”) y el

Tabla 3.1: Parámetros intrínsecos

Anchura de la imagen	640 pixels
Altura de la imagen	480 pixels
Distancia focal	6,28 mm
Anchura de la CCD	7,780 mm
Altura de la CCD	3,589 mm
Coordenada X del centro de la CCD	342,58 pixels
Coordenada Y del centro de la CCD	261,43 pixels
Distancia entre cámaras (b)	148,91 mm

Tabla 3.2: Parámetros del AG

Representación de la población	Coma flotante
Población	1000
Nº de padres	90 % de la población
Nº de hijos	90 % de la población
Probabilidad de sobrecruzamiento	70 %
Probabilidad de mutación	1 %
Máximo número de generaciones	50

Tabla 3.3: Resultados del AG

Parámetros	Altura (mm)	Cabeceo		Balanceo	
		(rad)	(°)	(rad)	(°)
Reales	1500,0	0,1700	9,74	-0,0900	-5,15
Estimados	1458,0	0,1691	9,69	-0,0933	-5,35
Varianza	1,34	$9 \cdot 10^{-11}$		$2 \cdot 10^{-7}$	
Error	42,0	0,0009	0,05	0,0033	0,20

Tabla 3.4: Más ejemplos de resultados del AG

Parámetros	Altura (mm)	Cabeceo		Guiñada	
		(rad)	(°)	(rad)	(°)
Reales	1500,0	0,1700	9,74	0,0900	5,16
Estimados	1484,3	0,1680	9,62	0,0811	4,65
Reales	1100,0	0,1700	9,74	-0,0900	-5,16
Estimados	1064,8	0,1682	9,64	-0,1001	-5,74

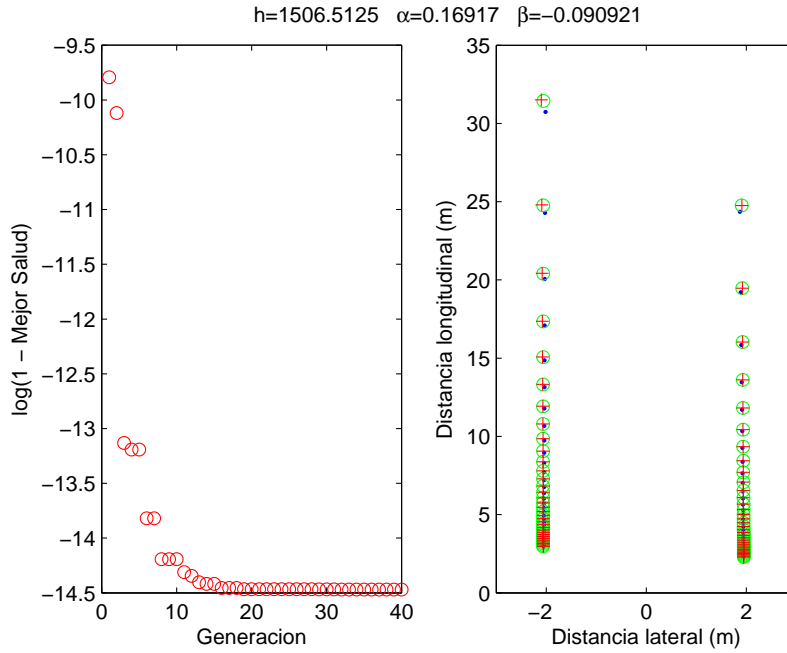
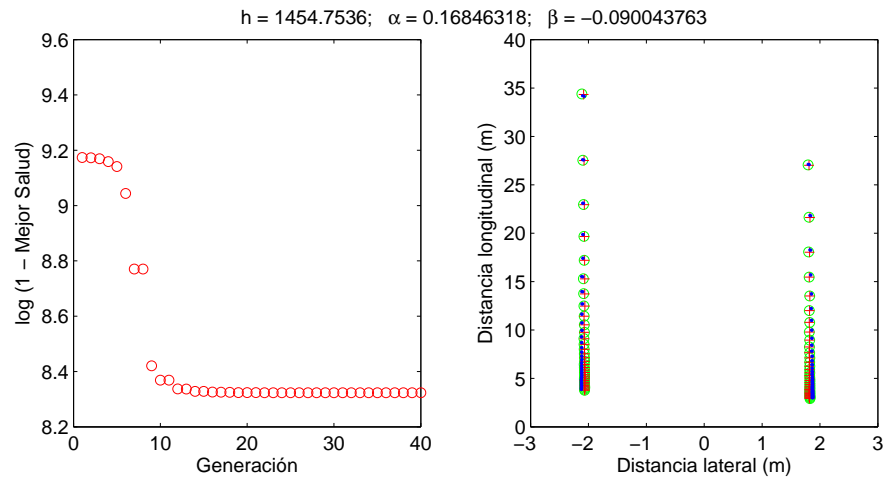


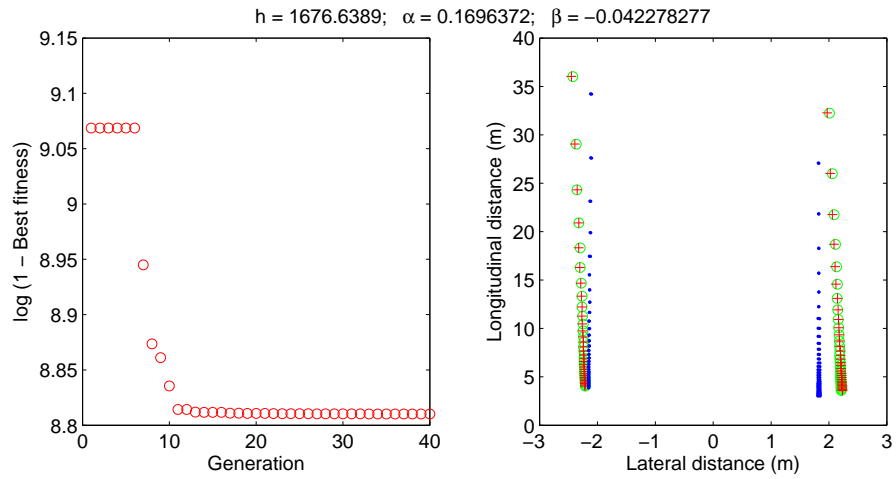
Figura 3.14: Resultados del AG; (+) puntos proyectados desde la cámara derecha; (o) puntos proyectados desde la cámara izquierda; (·) puntos proyectados según los parámetros reales

patrón proyectado desde la cámara izquierda (con “o”). También aparece el patrón proyectado utilizando los parámetros reales (con “·”), lo que representa la posición real del carril en la carretera. Esta gráfica permite comparar las posiciones estimadas de los objetos en la carretera (representadas por los círculos y las cruces), con las reales (representadas por puntos). Puede apreciarse que el error en este caso es de aproximadamente 1 metro a 35 metros de distancia al vehículo ($\approx 3\%$), y de unos 5 centímetros a 4 metros ($\approx 1.3\%$). La tabla 3.4 presenta más resultados con diferentes posiciones y orientaciones de las cámaras.

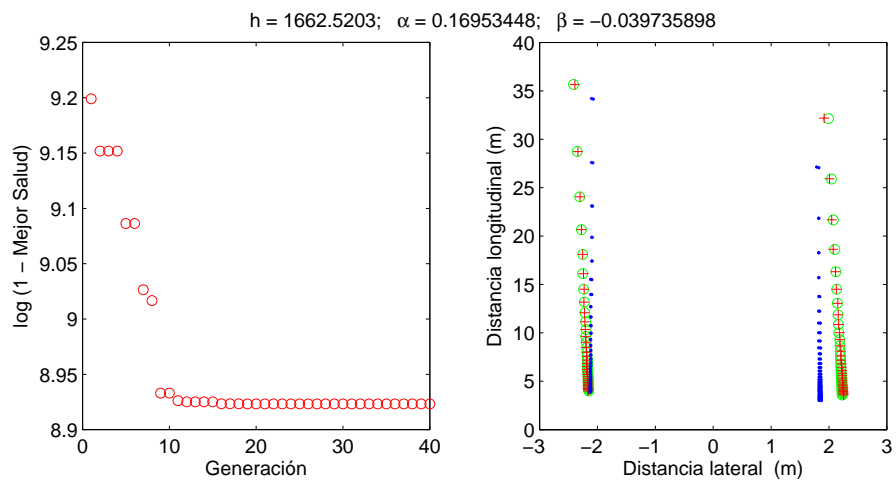
Para recrear un entorno más real, y así poder estudiar la sensibilidad del algoritmo a errores en la extracción del patrón de calibración, se han hecho pruebas añadiendo ruido gaussiano a las imágenes sintéticas. Los resultados muestran que el ruido afecta fundamentalmente a la estimación de la altura y el ángulo de balanceo, sin embargo la estimación del cabeceo apenas se ve afectada. La tabla 3.5 y la figura 3.15 muestran algunos resultados con varios niveles de ruido. Con un 5% de ruido, las estimaciones son igual de buenas que con la imagen limpia. Por encima del 10%, el error en la estimación sobre el plano de la carretera sube a 40 cm en posición a 4 m de distancia (10%), y de 5 m a 30 m del vehículo (16%). Esto se debe a que la detección de los bordes de las marcas viales empieza a fallar con ese nivel de ruido, y en consecuencia las líneas detectadas por la Transformada de Hough no se ajustan a la realidad.



(a) 5 % de ruido



(b) 10 % de ruido



(c) 20 % de ruido

Figura 3.15: Resultados con ruido gaussiano añadido a las imágenes sintéticas

Tabla 3.5: Resultados con ruido gaussiano añadido a las imágenes sintéticas.

Parámetros	Altura (mm)	Cabeceo (rad) (°)		Balanceo (rad) (°)	
Reales	1500,0	0,1700	9,74	-0,0900	-5,16
con 5 % de ruido	1454,9	0,1685	9,65	-0,0900	-5,16
con 10 % de ruido	1676,6	0,1696	9,72	-0,0423	-2,42
con 20 % de ruido	1662,5	0,1695	9,71	-0,0397	-2,28

3.3.2. Con Imágenes Reales

En el caso de imágenes reales, no se conocen a priori los parámetros de calibración reales, de modo que para evaluar el desempeño del algoritmo se ha ejecutado en 10 fotogramas consecutivos de una secuencia tomada en un tramo de carretera que cumple los requisitos exigidos de ser plano y recto. La figura 3.16 muestra los primeros fotogramas izquierdo y derecho de la secuencia, junto a las líneas detectadas a partir de las cuales se generará el patrón de calibración, mientras que la figura 3.17 presenta el resultado del algoritmo genético para el mismo caso.

La figura 3.18 muestra los parámetros estimados a lo largo de una secuencia de 10 fotogramas consecutivos. Al estar el coche en movimiento, resulta imposible aislar la influencia del balanceo y cabeceo provocado por la suspensión, por lo que no son del todo comparables los fotogramas entre sí, pero sí puede decirse que los resultados son coherentes entre sí, y acordes a la posición de montaje del sistema de visión.

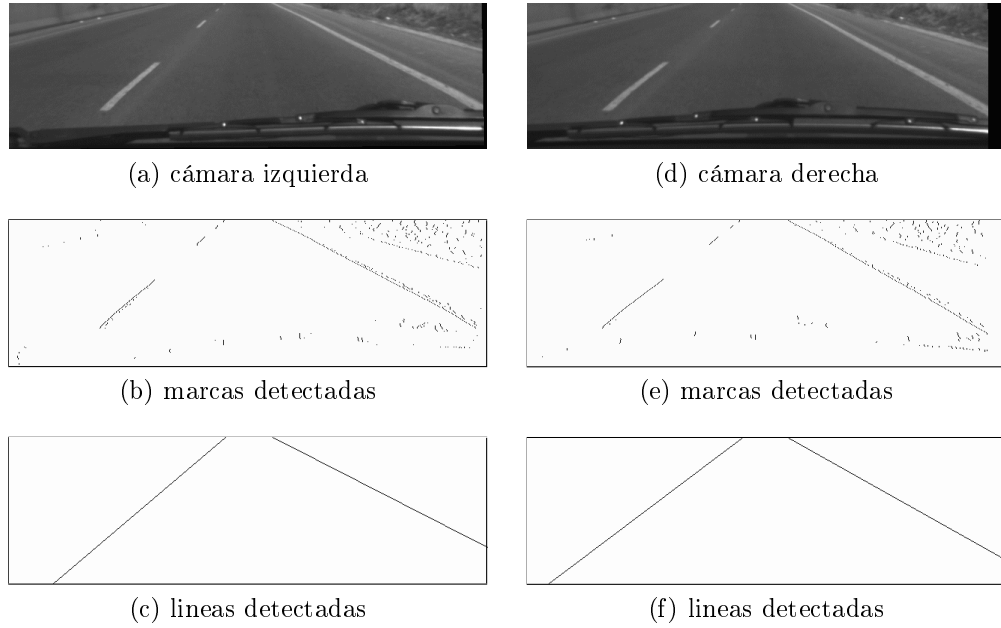


Figura 3.16: Detección de líneas y generación del patrón en el primer fotograma de la secuencia de imágenes reales.

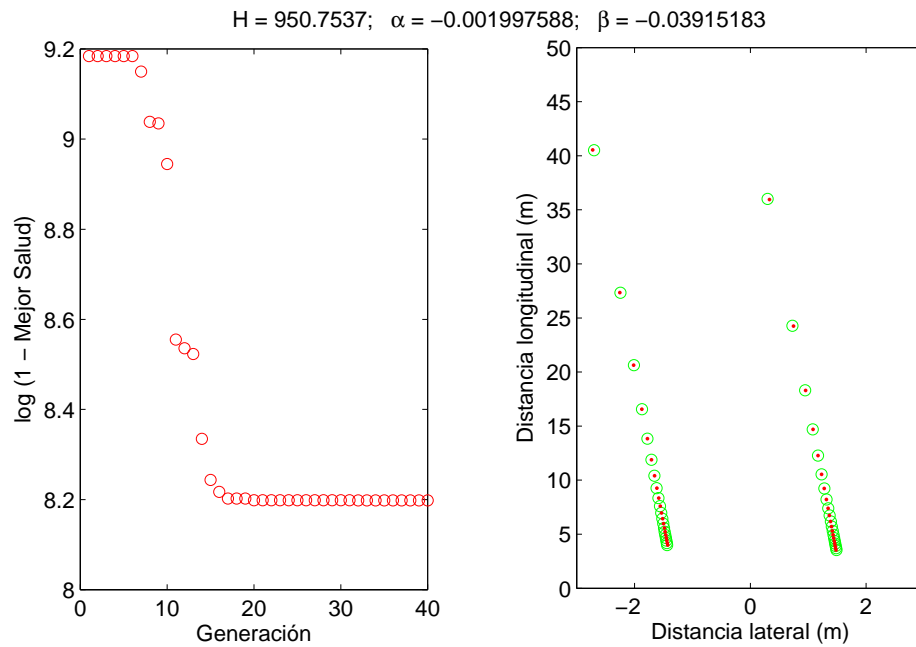
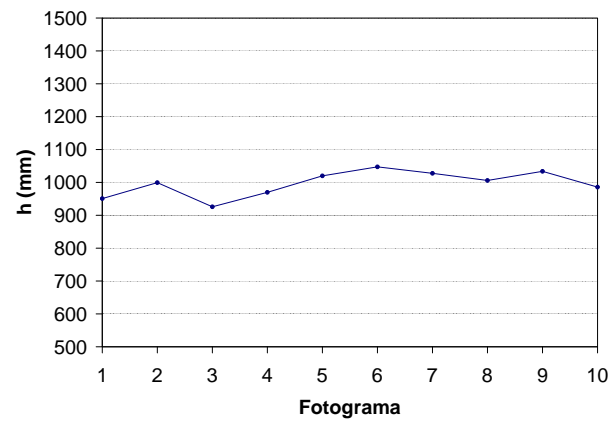
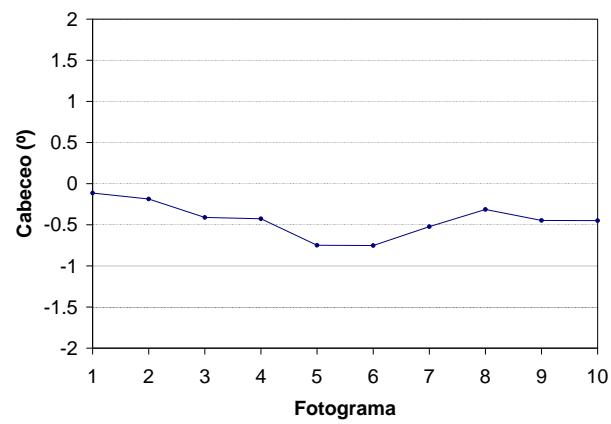


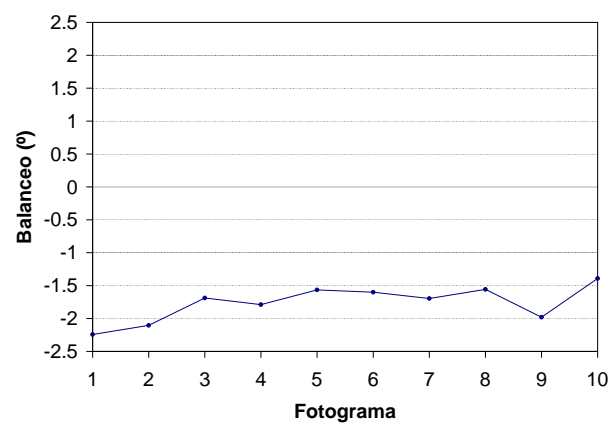
Figura 3.17: Resultado del Algoritmo Genético en el primer fotograma de la secuencia de imágenes reales.



(a) Evolución de la altura



(b) Evolución del ángulo de cabeceo



(c) Evolución del ángulo de balanceo

Figura 3.18: Parámetros extrínsecos a lo largo de una secuencia

Capítulo 4

Detección y seguimiento de carriles

4.1. Introducción

En este capítulo se describen los diferentes algoritmos que se han desarrollado para detectar y seguir las líneas viales. En general, todos estos algoritmos siguen el flujograma de la figura 4.1.

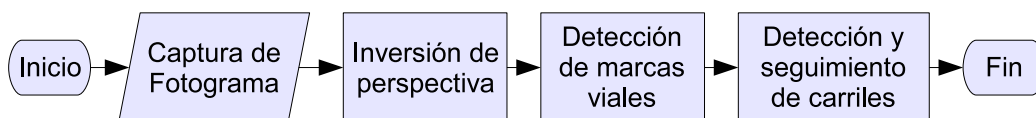


Figura 4.1: Organigrama General

El primer paso es la captura de la imagen mediante un sistema estéreo fijado al parabrisas, cuyas características se pueden consultar en el apéndice A. El sistema estéreo es necesario para la autocalibración, pero para la detección de carriles sólo se utiliza la cámara de la izquierda. El segundo paso es una transformación de perspectiva que se describe en el apartado 4.2. En el tercer paso se detectan las marcas viales longitudinales (apartado 4.3), y en el cuarto y último se ajusta a las marcas viales detectadas un modelo de carretera. Este paso es el que mayor evolución ha sufrido, y a él se dedican los apartados 4.4 y 4.5).

Cada nueva versión trata de solventar los inconvenientes de la anterior, de manera que, a grosso modo, se pueden distinguir tres etapas en la evolución de estos algoritmos:

1. *Detección de líneas viales rectas fotograma a fotograma:* Se utiliza primero un modelo de línea vial recta y después un modelo de carril recto. Se procesa cada fotograma de manera independiente.
2. *Detección de carriles rectos con el filtro de Kalman:* Se define un modelo de carril recto, y se incorpora una técnica de seguimiento.



(a) Flechas impresas sobre el pavimento.



(b) Sombras producidas por los guardarrailes.

Figura 4.2: Dificultades en la detección de líneas viales (I).

3. *Detección de carriles curvos con el filtro de partículas*: Se mejora el modelo de carril y se utiliza otra técnica de seguimiento más robusta.

Recordemos que el objetivo que corresponde a la detección y seguimiento de carriles es poder reconocer e interpretar la mayor cantidad de información relevante de la carretera, siendo lo ideal llegar a detectar el tipo de carretera por el que se circula (autovía, carretera secundaria, etc.). Esto exige, además de la capacidad para clasificar las líneas viales que se explica en el siguiente capítulo, detectar de manera precisa y robusta la posición y el perfil de todos los carriles de la carretera.

La detección de líneas viales presenta las siguientes dificultades:

- *Ruido estructurado*, entendido éste como el producido por algún patrón regular. Por ejemplo, la presencia de otros objetos en la imagen que, al proyectarse sobre la imagen de perspectiva inversa, dan lugar a formas similares a marcas viales, puede confundir fácilmente al algoritmo. Las estructuras más frecuentes que provocan confusión son:
 - Otras marcas viales sobre el pavimento, como flechas o letras (figura 4.2a).
 - Sombras de forma alargada y paralelas a la carretera, producidas por guardarrailes u otros objetos (figura 4.2b).



(a) Vehículos ocultando parcialmente las líneas viales.



(b) Líneas borradas mediante fresado.

Figura 4.3: Dificultades en la detección de líneas viales (II).

- *Oclusiones*: debidas a sombras fuertes o a otros vehículos (figura 4.3a).
- *Defectos de borrado*: Determinadas técnicas de borrado de las marcas viales no consiguen eliminarlas completamente. Por ejemplo, el borrado mediante fresas o discos dentados suele dejar un surco visible (figura 4.3b). También, en el caso de borrado con pintura o lechada asfáltica, cuando la luz incide con un determinado ángulo es reflejada por la pintura superpuesta de tal manera que las líneas borradas llegan a resultar tan visibles o más como las marcas verdaderas, y con apariencia muy similar.
- *Defectos de pintura*: Cuando las marcas viales están deterioradas o parcialmente borradas, su bajo contraste con el resto de la calzada las hace más difíciles de detectar (figura 4.4a).
- *Ruido no estructurado*: La presencia de suciedad, polvo, parches de asfalto, grietas, y en general cualquier falta de homogeneidad en la calzada genera ruido que es necesario filtrar durante el procesado (figura 4.4b).
- *Condiciones climáticas y ambientales*:
 - *Brillos*. Por ejemplo, cuando el Sol está próximo al horizonte, y en especial cuando aparece de frente al vehículo, se producen reflejos en la textura granulada del asfalto que confieren a la calzada un brillo que enmascara las marcas viales (figura 4.4c).



(a) Líneas viales deterioradas



(b) Parches de asfalto



(c) Brillos en el asfalto.



(d) Reflejos provocados por la lluvia.



(e) Saturación de la imagen a la salida de un túnel.

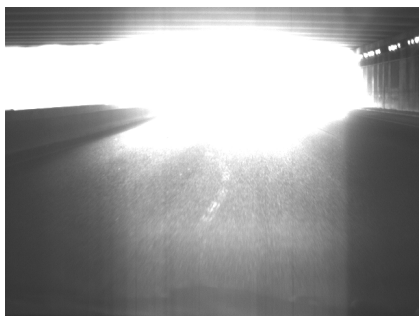


Figura 4.4: Dificultades en la detección de líneas viales (y III).

- *Lluvia*. El problema que la lluvia ocasiona no se debe a la iluminación, de hecho, las nubes producen iluminación difusa y suprimen el problema de los reflejos, con lo que suele mejorar la eficacia de los algoritmos de visión en exteriores. El principal inconveniente lo produce el aumento de la reflectividad del pavimento mojado (en especial cuando éste es poco poroso), y que como en el caso anterior, puede llegar a enmascarar las marcas viales (figura 4.4d).
- *Cambios bruscos de iluminación*. El caso más típico son las salidas de túneles, en los que la velocidad de cambio de iluminación supera el rango dinámico del sistema de visión y la imagen se satura durante 2-3 fotogramas (figura 4.4e).

4.2. Perspectiva inversa

Una vez capturada la imagen, hay que decidir sobre la conveniencia de realizar una inversión de perspectiva (figura 4.5). Esta operación consiste en eliminar el efecto de la perspectiva en la carretera, obteniendo a partir de la imagen capturada (figura 4.6a), una nueva imagen de la carretera a vista de pájaro (figura 4.6b), lo cual permite trabajar directamente sobre el sistema de coordenadas de la carretera, en lugar del de la cámara.

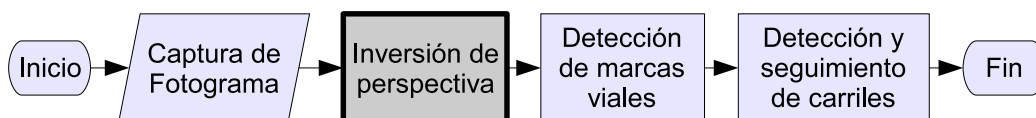


Figura 4.5: Etapa 2: Inversión de perspectiva

Trabajar en el sistema de coordenadas de la cámara (figura 4.6a), es decir, directamente sobre la imagen capturada, tiene la ventaja de que la imagen de partida no ha sufrido pérdida de resolución ni calidad, ya que no ha sido transformada. Además disponemos de campo de visión amplio, similar al del conductor.

Por otra parte, trabajar en el sistema de coordenadas de la carretera (figura 4.6b) requiere que la imagen sea transformada, con la consiguiente pérdida de resolución, calidad y campo de visión, y es necesario que la carretera sea plana al menos en un entorno de unos 30-40m. Sin embargo, al reducir su tamaño y limitar el campo de visión estrictamente a la zona de interés, permite que el procesado posterior sea más rápido. También facilita dicho procesado, ya que las líneas viales tienen una apariencia homogénea en toda la imagen, es decir, su anchura es la misma independientemente del punto de la imagen al que se mire, lo cual no sucede en la imagen capturada por la cámara, en la que las líneas se estrechan a medida que se acercan al horizonte. Además, se está trabajando ya en un sistema de coordenadas relativo a la carretera, y no a la cámara. Esto evita que el modelo de carretera tenga que ser traducido de un sistema de coordenadas a otro, y reduce en gran medida las modificaciones que serían necesarias en el algoritmo en el caso de cambiar el sistema de visión.

4.2.1. Implementación de la perspectiva inversa

La transformación de perspectiva se realiza mediante matrices homogéneas, con el mismo procedimiento que se describe en el apartado 3.6.

La nueva imagen deberá tener, por una parte, un tamaño y resolución suficientemente elevados como para conservar los detalles de la imagen original y observar un tramo razonable de carretera, y por otra parte, un tamaño suficientemente pequeño como para no requerir un tiempo de procesamiento largo.

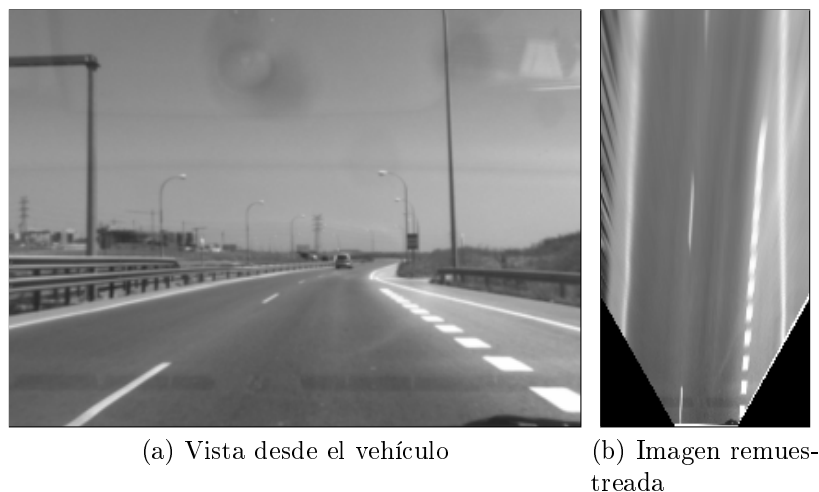


Figura 4.6: Transformación de perspectiva

En las primeras versiones se utilizó un tamaño de imagen de 128×256 píxeles con una resolución de 10 cm/píxel, representando así un tramo de carretera de $25,6 \times 12,8$ m cuyo origen se sitúa a 5 m del vehículo, que es el punto a partir del cual la carretera comienza a ser visible desde las cámaras del IVVI. Por lo tanto el tramo remuestreado comprende desde los 5 m hasta los 30,6 m. A medida que las etapas posteriores del algoritmo fueron mejorando, este tamaño se amplió a 140×350 píxeles (de 5 a 40 m). Si se estima que el tiempo de reacción de un humano está en torno a un segundo, la distancia de análisis no deberá ser inferior a 33 m (distancia recorrida en ese tiempo a 120 km/h), de modo que 40 m se considera una distancia razonable.

4.3. Detección de marcas viales

Una vez realizada la transformación de perspectiva, la tercera etapa (figura 4.7) consiste en resaltar en la imagen aquellos píxeles que pertenecen a una marca vial. Las características que identifican estas marcas son dos:

- su *color* o *nivel de gris*: las marcas aparecen como bandas de color claro sobre un fondo oscuro.
- su *dirección*: en la imagen en perspectiva inversa las marcas aparecen como bandas verticales, ya que se trata de un entorno próximo al vehículo¹.

¹Según la *Norma 3.1-IC. Trazado* (BOE, 2000), en autopistas, autovías, vías rápidas y carreteras C-100, el radio mínimo de curvatura es de 250 m. Esto significa que a 40 m del vehículo, la orientación de las marcas viales no puede exceder de $9,1^\circ$

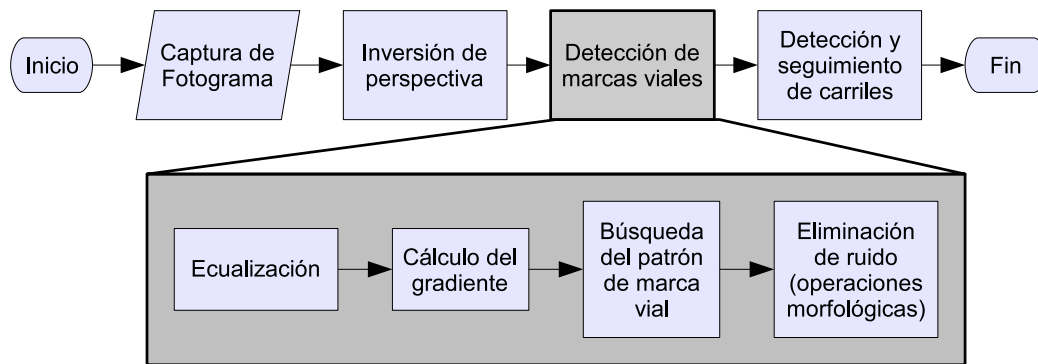


Figura 4.7: Etapa 3: Detección de marcas viales

Por tanto, la búsqueda de píxeles que pertenecen a marcas viales consiste en localizar transiciones oscuro-claro-oscuro en cada fila de la imagen. Estas transiciones son bordes verticales que corresponden a puntos de máximo gradiente en dirección horizontal. El algoritmo de detección de marcas viales se compone de las siguientes cuatro fases (figura 4.7):

1. Ecualización de la imagen en perspectiva inversa.
2. Cálculo del gradiente en dirección horizontal.
3. Búsqueda de parejas de máximos locales en el gradiente.
4. Eliminación de ruido.

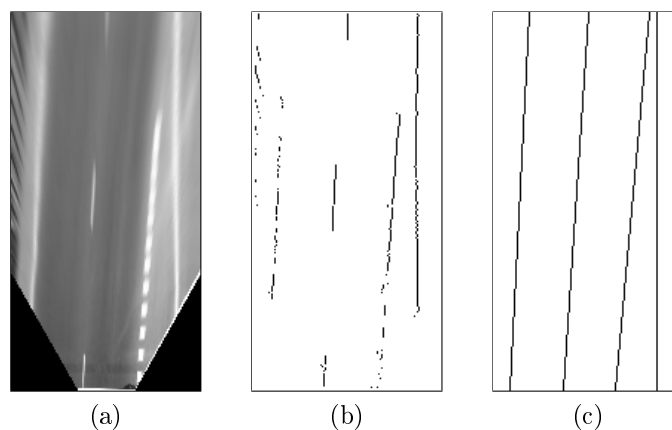


Figura 4.8: (a) imagen remuestreada, a vista de pájaro; (b) candidatos a marca vial; (c) líneas detectadas por la Transformada de Hough.

4.3.1. Ecualización de la imagen

Para independizar el procesamiento de la iluminación, se realiza primero una ecualización consistente en ampliar el histograma para ocupar todo el rango de valores y aumentar el contraste. La figura 4.9 muestra el resultado.

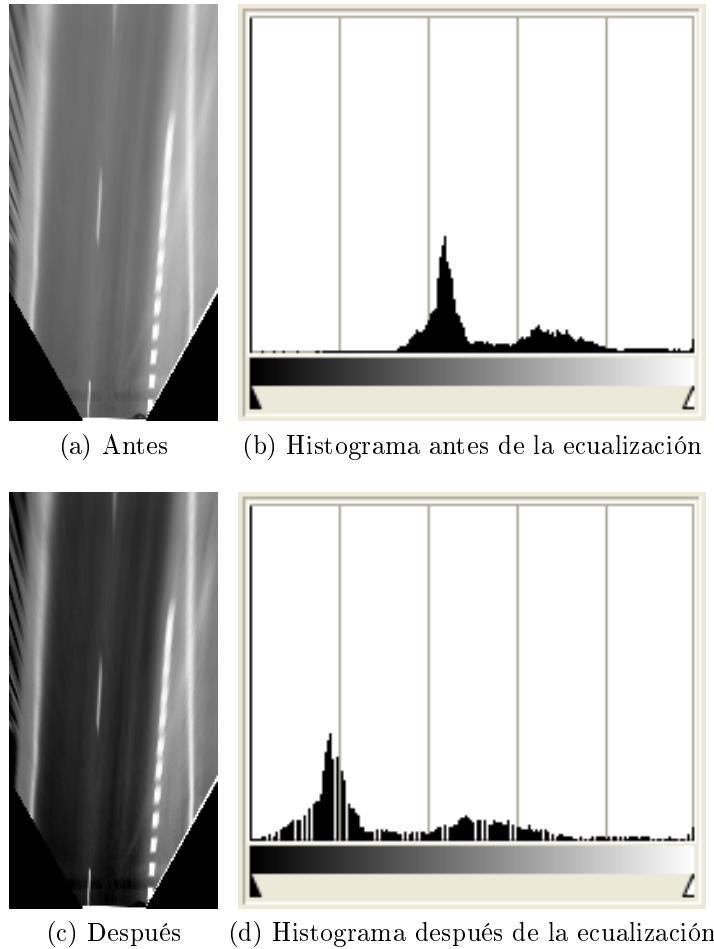


Figura 4.9: Ajuste del histograma de la imagen.

4.3.2. Cálculo del gradiente

Se han estudiado varias alternativas para calcular el gradiente horizontal de una imagen:

1. **Filtro lineal con máscara 1×3 :** Se filtra la imagen con la máscara de la matriz 4.1. La ventaja de este filtro es la rapidez.

$$\begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \quad (4.1)$$

2. **Filtro lineal con máscara 3×3 :** Como el anterior, pero con una máscara más amplia (matriz 4.2), que lo hace menos sensible al ruido.

-1	0	1
-2	0	2
-1	0	1

(4.2)

3. **Filtro basado en Canny:** Se trata de filtro espacial que utiliza la primera etapa del filtro de Canny para calcular la orientación del borde. Este filtro permite restringir los bordes detectados a aquellos que tengan una determinada orientación (en este caso, bordes verticales), y devuelve bordes de un píxel de anchura.

La figura 4.10 muestra una comparativa de los tres filtros. Las máscaras 1×3 (figura 4.10b) y 3×3 (figura 4.10c) no presentan grandes diferencias en los resultados, sin embargo, el filtro basado en Canny (figura 4.10d) devuelve una imagen mucho más nítida, y a diferencia de los anteriores, permite obtener bordes de un sólo píxel de anchura.

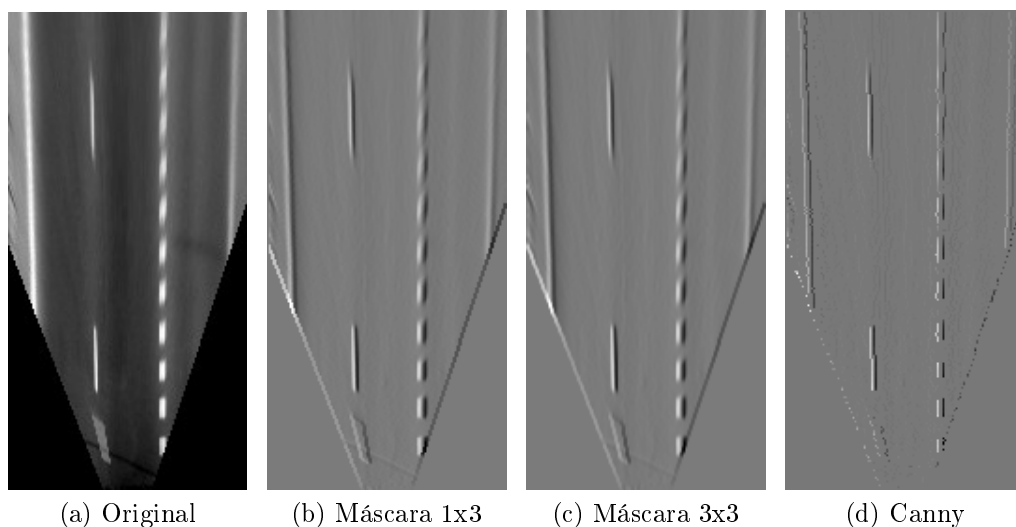


Figura 4.10: Gradientes

4.3.3. Búsqueda del patrón de marca vial

El siguiente paso consiste en buscar la huella que una marca vial genera en la imagen del gradiente. Como se introdujo al principio del capítulo, las marcas viales son bandas verticales de color claro sobre un fondo oscuro, de modo que en el gradiente horizontal deben aparecer dos bordes próximos entre sí, con sentidos opuestos.

La búsqueda de marcas viales se hace de forma muy similar a que utiliza el algoritmo de autocalibración (apartado 3.2.2). El algoritmo busca en

cada fila de la imagen del gradiente un patrón formado por dos picos, el primero de signo positivo y el segundo de signo negativo, que corresponden, respectivamente, a los bordes izquierdo (transición oscuro-claro) y derecho (transición claro-oscuro). Ambos picos de gradiente deben estar espaciados una distancia igual al ancho de la marca vial. En la práctica se fija un rango que va de 1 a 6 píxeles. Cuando se localiza este patrón, se marca el píxel central (equidistante a los dos picos) como perteneciente a una marca vial (figura 4.11).

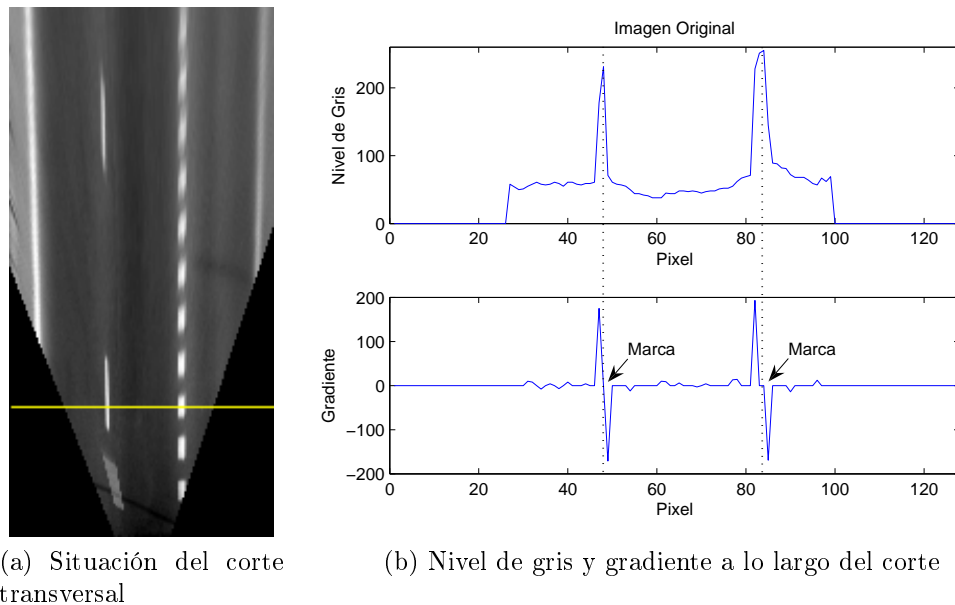


Figura 4.11: Perfil de un corte transversal en la imagen original y en el gradiente.

4.3.4. Filtrado de ruido con operaciones morfológicas

Cuando las marcas viales están deterioradas y despintadas por el uso, o cuando no presentan un contraste suficiente con el entorno, el proceso de detección devuelve una imagen en la que las marcas viales aparecen interrumpidas. De la misma manera el ruido de la imagen puede dar lugar a la detección de falsos positivos.

El objetivo de las siguientes etapas de procesamiento es limitar la presencia tanto de falsos positivos como de falsos negativos producidos por efecto del ruido de la imagen:

- **Transformación “Acierta o Falla” (“*Hit or Miss*”):** Para recuperar la presencia de falsos negativos, se realiza una transformación “Acierta o Falla”, cuyo efecto es el de “estirar” los bordes verticales, hacia arriba y hacia abajo (figura 4.12b).

Para ello se utilizan las siguientes seis máscaras, en las que el elemento con fondo gris indica el píxel central:

X	1	X
0	1	0
0	0	0

1	0	0
0	1	0
0	0	0

0	0	1
0	1	0
0	0	0

0	0	0
0	1	0
X	1	X

0	0	0
0	1	0
1	0	0

0	0	0
0	1	0
0	0	1

- **Eliminación de píxeles aislados:** Esta última etapa pretende eliminar los falsos positivos, excluyendo de la detección aquellos elementos que no tienen un área mínima, en este caso de 2 píxeles (figura 4.12c).

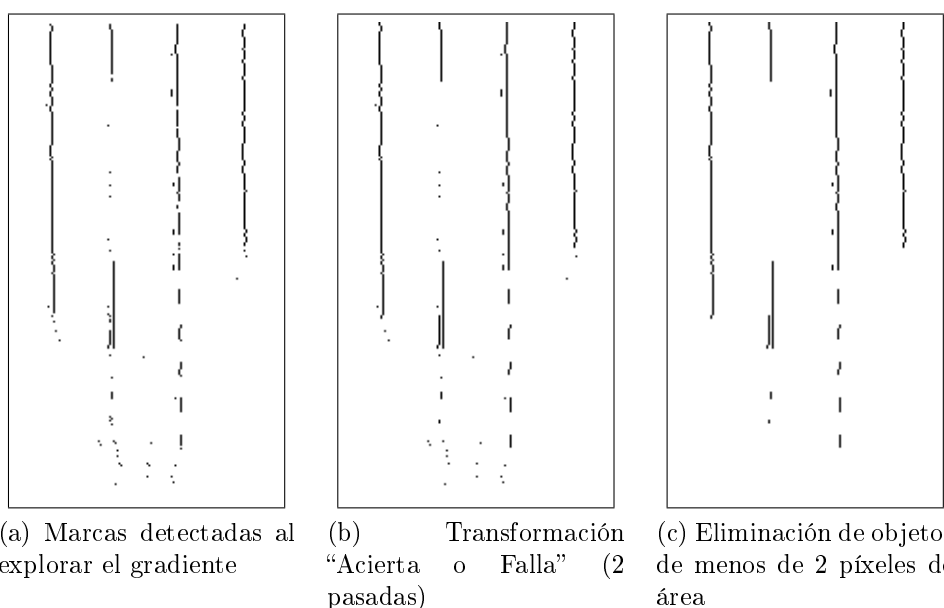


Figura 4.12: Etapas de la detección de marcas.

Esta etapa de filtrado se añadió durante la implementación del algoritmo de seguimiento para carriles curvos, descrito en el apartado 4.4. Este algoritmo hace uso de la transformada distancia (figura 4.30), que se ve muy influenciada por la presencia de píxeles aislados provocados por el ruido. El algoritmo de seguimiento de carriles rectos (apartado 4.5) no necesita este filtrado, ya que se basa en la transformada de Hough, técnica que fue diseñada para trabajar en entornos ruidosos.

4.4. Modelo de carretera recta

A partir de este apartado se describen los experimentos realizados durante el desarrollo de la última etapa del algoritmo (figura 4.13).

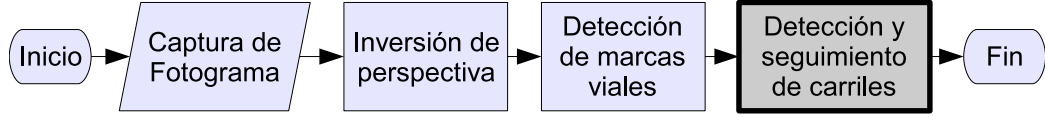


Figura 4.13: Etapa 4: Detección y seguimiento de carriles.

La primera versión del algoritmo consideraba la carretera como recta y plana. Estas dos simplificaciones pueden considerarse válidas en carreteras interurbanas en un entorno próximo al vehículo (hasta unos 40 metros).

La ventaja de este modelo es la sencillez y simplicidad que permite representar un carril con sólo dos parámetros. Este espacio de búsqueda tan reducido puede explorarse de forma exhaustiva mediante técnicas como la transformada de Hough (Leavers, 1992). Ésta técnica fue diseñada para la detección de trayectorias de partículas subatómicas en cámaras de niebla (Hough, 1959), es decir, con información parcial y en entornos con gran presencia de ruido. Se trata de una técnica que analiza la imagen de forma global, lo que le da la capacidad para localizar patrones incluso en situaciones de oclusión parcial o de gran nivel de ruido.

4.4.1. Detección de líneas viales con la Transformada de Hough

La primera implementación del algoritmo ajustaba un modelo recto de línea vial (figura 4.14). Es decir, localizaba de forma independiente las líneas viales presentes en la imagen, sin agruparlas en carriles.

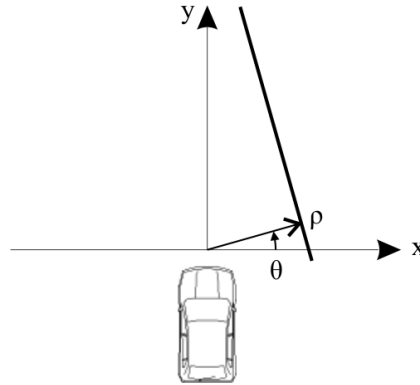


Figura 4.14: Modelo de línea vial con parametrización ρ - θ ($\rho = x \cos \theta + y \sin \theta$)

Dado que el entorno es muy ruidoso, una técnica que utilice información global, como la transformada de Hough, parece adecuada a este caso.

La parametrización de las rectas que utilizaremos aquí es la denominada parametrización ρ - θ ($\rho = x \cos \theta + y \sin \theta$) (figura 4.14), habitualmente utilizada con esta transformada, ya que evita la indeterminación a que da lugar la parametrización clásica basada en la pendiente y la ordenada en el origen ($y = mx + b$), para el caso de pendientes infinitas.

Se ha demostrado en (Palmer y otros, 1993b) que el mejor perfil de la función de votación, también denominada *función núcleo* de Hough es una gaussiana (figura 4.15a). En (Palmer y otros, 1993a) se compara el rendimiento de dicha función con el de la función “sombrero” (figura 4.15b), concluyendo que utilizar la gaussiana como función núcleo da lugar a picos más abruptos en el acumulador de Hough que son más fáciles de detectar. El inconveniente es que requiere operar en coma flotante, lo cual resulta caro computacionalmente. Aquí se ha decidido utilizar una aproximación (figura 4.15c), que ofrece más rapidez y una precisión suficiente. Esta función núcleo se aplica en la dirección del parámetro ρ .

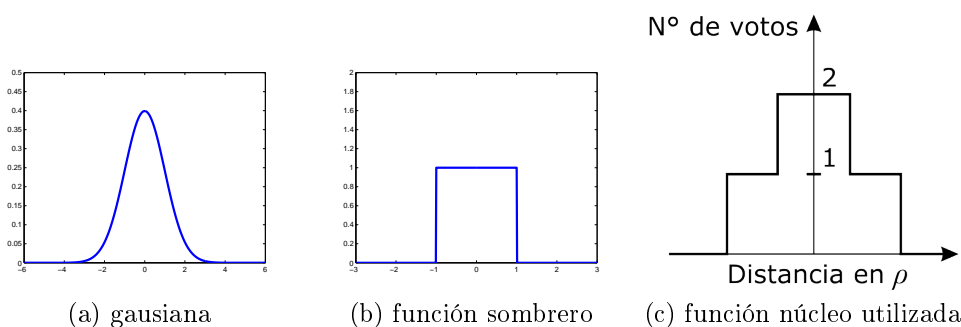


Figura 4.15: Funciones de votación para la transformada de Hough.

Para mejorar la respuesta de la Transformada de Hough, en lugar de tomar directamente los parámetros de la línea más votada, correspondientes a un máximo local del acumulador de Hough, éstos se calculan como media ponderada de las líneas que quedan englobadas en un entorno cercano al máximo, dentro de la matriz del acumulador. Como pesos para la ponderación, se toma el número de votos que recibe cada línea. Esta mejora incrementa ligeramente la precisión en los parámetros de las líneas detectadas.

De nuevo, puesto que las líneas son prácticamente verticales, el acumulador de Hough puede acotarse en ángulo. Esto evita interferencias con otros objetos de la imagen y acelera el cálculo. El rango permitido en la inclinación de las líneas va de -15 a 15 grados sexagesimales.

La figura 4.16 muestra un ejemplo. A la izquierda (figura 4.16a) aparece la vista desde la cámara. El recuadro blanco delimita la zona procesada, dentro de la cual se muestran las líneas detectadas en color rojo, verde y azul. En el centro (figura 4.16b) tenemos la imagen en perspectiva inversa, sobre la cual

aparecen en amarillo las marcas viales detectadas, y en el resto de colores, de nuevo las líneas encontradas. A la derecha (figura 4.16c), está representada la matriz del acumulador de Hough, con los máximos locales enmarcados en círculos.

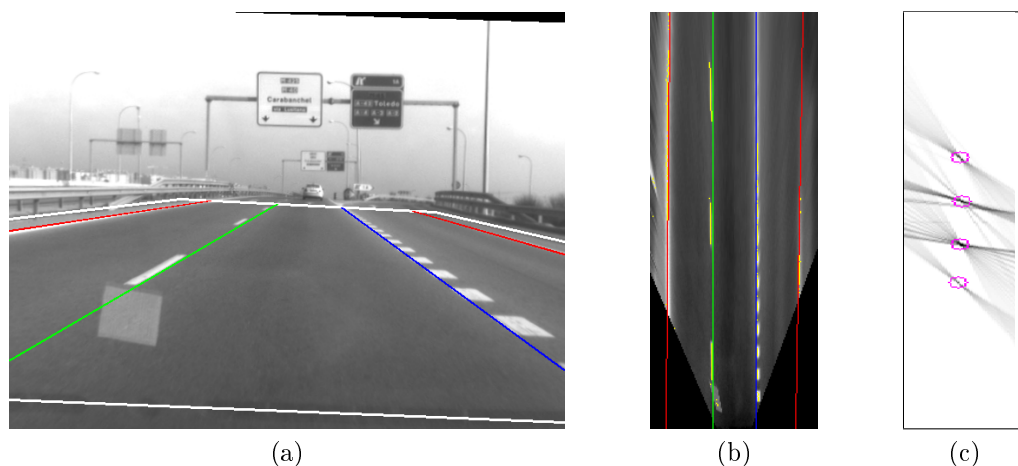


Figura 4.16: Detección de líneas viales con la transformada de Hough; (a) vista desde la cámara; (b) perspectiva inversa; (c) detalle del acumulador de Hough

4.4.2. Detección de líneas viales en Regiones de Interés (RDIs)

En el caso anterior se detectan todas las posibles líneas viales, sin discriminar si pertenecen a un carril o no, lo que hace al algoritmo muy vulnerable al ruido estructurado. Con frecuencia aparecen falsos positivos como líneas que se cruzan (figura 4.19), o líneas que corresponden a otras marcas viales como por ejemplo, flechas (figura 4.18).

Para limitar la presencia de falsos positivos, debidos fundamentalmente a que se trabaja con un simple modelo de línea vial, se define ahora un modelo de carril como una combinación de varias líneas viales paralelas que están separadas entre sí aproximadamente la misma distancia, el ancho del carril. Este modelo incluye hasta cuatro líneas viales para las cuales se delimitan en el acumulador de Hough cuatro regiones de interés: dos correspondientes a los bordes izquierdo y derecho, más otras dos regiones correspondientes a los bordes de los carriles adyacentes (figura 4.17).

En cada región de interés se extrae la mejor línea vial utilizando la Transformada de Hough, y se considera válida cuando ha obtenido una puntuación suficiente. Primero se detectan los bordes del carril por el que circula el vehículo. Esos bordes se clasifican en tres clases, tal como se explica en el capítulo 5:

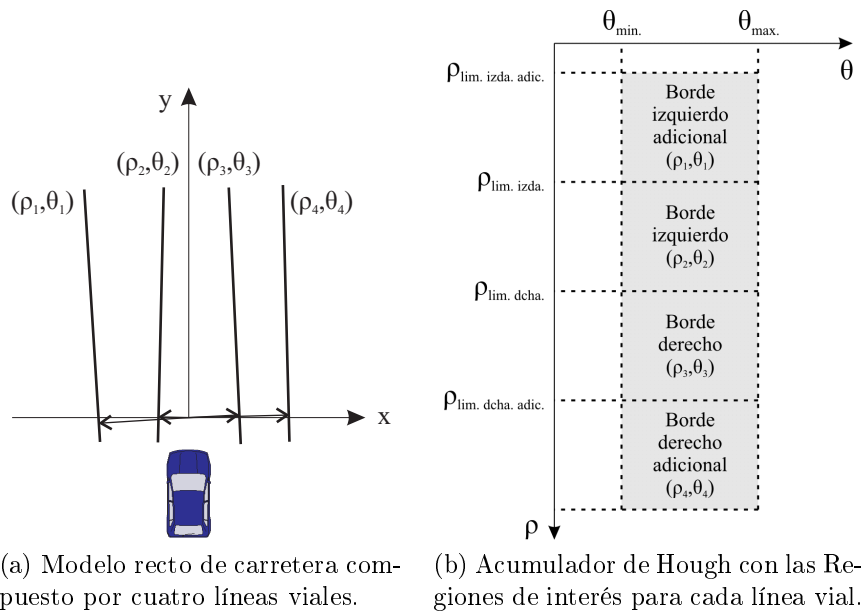


Figura 4.17: Modelo de carretera utilizando cuatro regiones de interés.

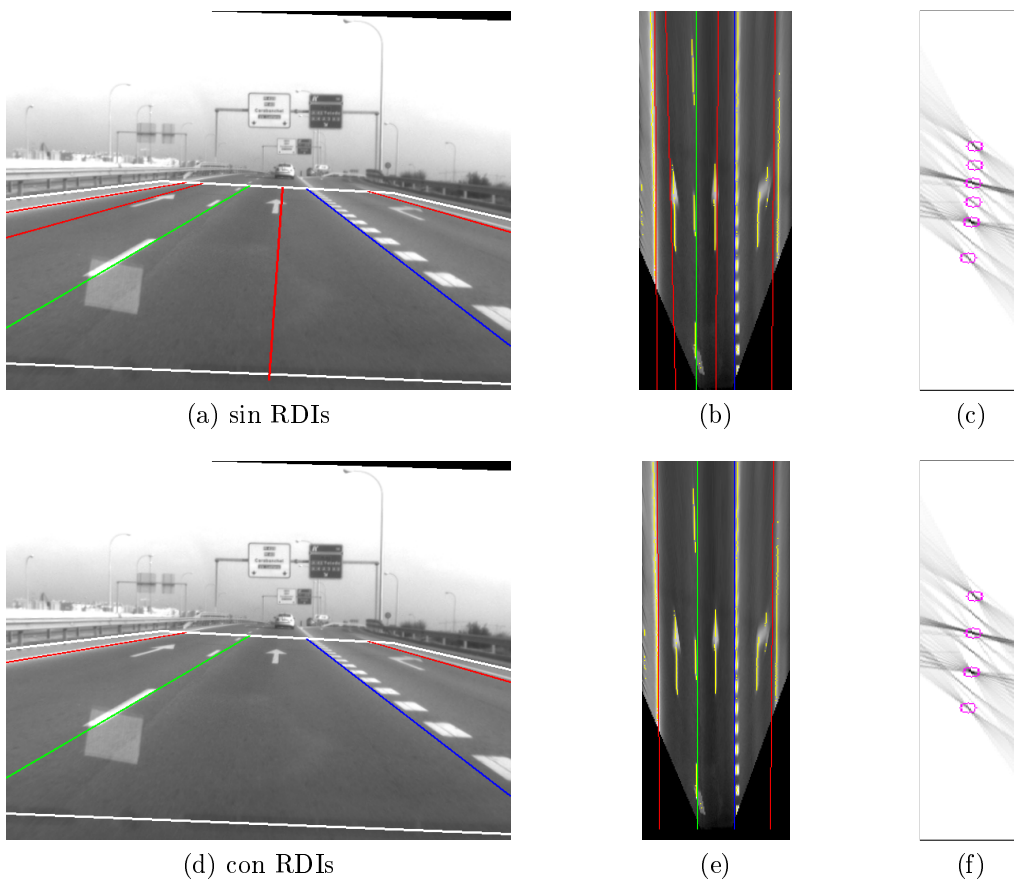


Figura 4.18: Detección de líneas viales con y sin Regiones de Interés

- línea *continua*,
- línea discontinua de separación de carriles normales, a la que nos referiremos como línea *discontinua*,
- línea discontinua de separación de carriles de entrada/salida, a la que nos referiremos como línea de *incorporación*.

En el caso de que un borde sea clasificado como línea *discontinua* o de *incorporación*, entonces, y sólo entonces, se explora la región correspondiente al borde del carril adyacente. De esta manera se limita la sensibilidad al ruido estructurado, ya que sólo se buscan una línea en cada región. La figura 4.18 presenta un ejemplo de cómo este mecanismo evita la detección de flechas.

Además, para evitar el problema de la detección de falsos positivos que se cruzan con otras líneas, se ha implementado un mecanismo para obtener la dirección mayoritaria de todas las líneas detectadas, que consiste en proyectar el acumulador de Hough sobre el eje del parámetro angular² (figura 4.19f). Sobre esta proyección se calcula el valor máximo, correspondiente a la orientación mayoritaria de las líneas de la imagen, y se delimita un entorno de 3 grados alrededor de éste. Fuera de este entorno se considera que las detecciones no son válidas. La figura 4.19 muestra un ejemplo en el que este sistema evita dos falsos positivos.

4.4.3. Seguimiento de carriles con el filtro de Kalman

El algoritmo del apartado anterior considera regiones de interés fijas en el tiempo. Para mejorar el rendimiento y la precisión del algoritmo, se implementó un filtro de Kalman que permitiese controlar la posición y el tamaño de las regiones de interés, a lo largo del tiempo.

En esta etapa, se define un nuevo modelo de carril que se compone de tres parámetros (figura 4.20):

- d es la distancia al eje del carril,
- θ es el ángulo que forman los ejes del vehículo y del carril,
- y W es el ancho del carril.

Sobre este modelo de carril, se construye un modelo de carretera (figura 4.21) que incluye los tres tipos de líneas viales considerados (*continua*, *discontinua*, y de *incorporación*) y el número de carriles que se calcula de forma automática a partir de la información de los tipos de líneas viales. Cabe recordar aquí que, mientras los modelos de carriles han sido ampliamente tratados en este área de la investigación, el número de trabajos dedicados

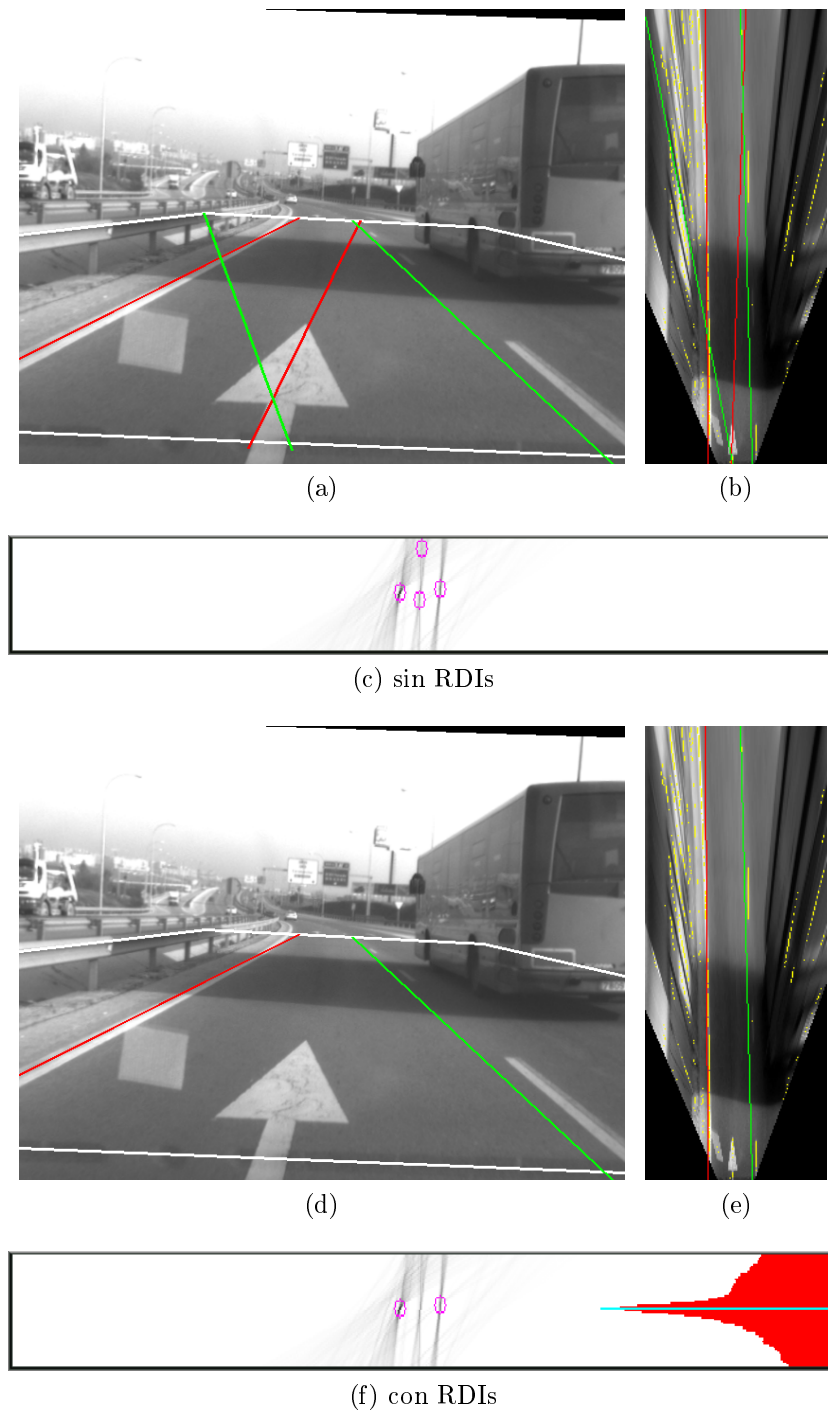


Figura 4.19: Detección impidiendo cruces entre líneas; (arriba) sin Regiones de Interés; (abajo) con Regiones de Interés y detección de la orientación mayoritaria de las líneas; (c) y (f) muestran los acumuladores de Hough rotados 90° respecto a las figuras anteriores; (f) presenta además la proyección sobre el eje angular a la derecha, que corresponde a la orientación mayoritaria de las líneas.

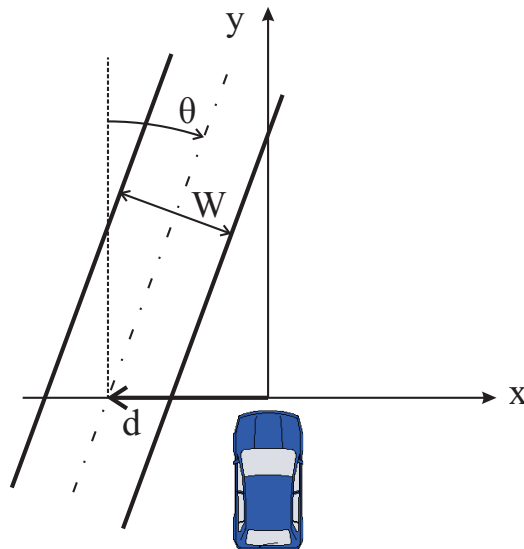


Figura 4.20: Modelo de carril recto

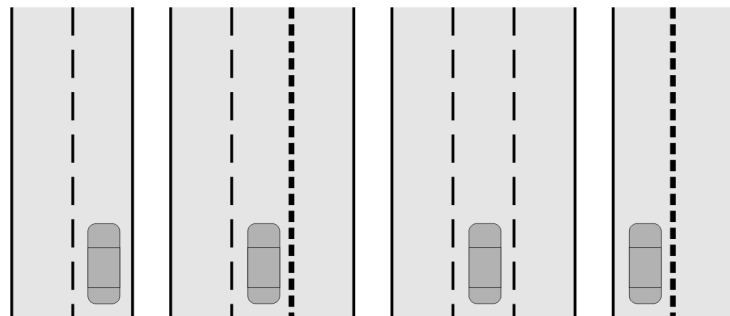


Figura 4.21: Modelo de carretera

a modelos de carretera más amplios, apoyados en la interpretación de los elementos detectados en la imagen, es considerablemente inferior.

Para el filtro de Kalman es necesario definir las ecuaciones de estado y de medida. Como primera aproximación, se ha considerado un vector de estado x_t compuesto por las siguientes cinco variables:

- la posición lateral (d),
- la velocidad de desplazamiento lateral (\dot{d}),
- la orientación del vehículo respecto al eje del carril (ϑ),
- su velocidad angular ($\dot{\vartheta}$),
- y el ancho de carril (W).

El vector de medida z_t se compone de las tres variables que definen el modelo de carril de la figura 4.20: d , θ y W .

En la ecuación de estado, se han considerado que las variables d , θ y W evolucionan de forma independiente. Por lo tanto, el modelo para el filtro de Kalman queda expresado por la ecuación 4.3, donde Δt es el tiempo transcurrido entre la captura de dos fotogramas consecutivos, F y H son las matrices de estado y de medida, que se asumen como constantes, y los vectores v_t y w_t representan los ruidos del proceso y de la medida, respectivamente:

$$\left\{ \begin{array}{l} \underbrace{\begin{bmatrix} \vartheta \\ \dot{\vartheta} \\ d \\ \dot{d} \\ W \end{bmatrix}}_{x_{t+1}} = \underbrace{\begin{bmatrix} 1 & \Delta t & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & \Delta t & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_F \cdot \underbrace{\begin{bmatrix} \vartheta \\ \dot{\vartheta} \\ d \\ \dot{d} \\ W \end{bmatrix}}_{x_t} + v_t \\ \\ \underbrace{\begin{bmatrix} \vartheta \\ d \\ W \end{bmatrix}}_{z_{t+1}} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}}_H \cdot \underbrace{\begin{bmatrix} \vartheta \\ \dot{\vartheta} \\ d \\ \dot{d} \\ W \end{bmatrix}}_{x_t} + w_t \end{array} \right. \quad (4.3)$$

Delimitación de Regiones de Interés en el acumulador de Hough

De nuevo, debido a que las líneas viales son prácticamente verticales en la imagen en perspectiva inversa, el parámetro ϑ se restringe al intervalo

²Esta operación consiste en sumar, para cada columna, el valor de todas las celdas de la matriz del acumulador

$[-15^\circ, +15^\circ]$. Una vez que el acumulador de Hough se ha calculado, la búsqueda de máximos locales se limita a ciertas regiones de interés. Estas regiones están centradas en la estimación dada por el filtro de Kalman, y su tamaño se calcula a partir de las varianzas de la posición lateral (ρ) y de la orientación del vehículo (ϑ).

En el primer fotograma, las regiones de interés se inicializan con el mismo tamaño que en el apartado 4.4.2, es decir, dividiendo todo el acumulador en cuatro zonas. En las dos zonas centrales, correspondientes a los bordes izquierdo y derecho del carril por el que circula el vehículo, se localizan los máximos locales del acumulador. Si se encuentran varias líneas en una región, el algoritmo intenta emparejarlas con otras líneas detectadas en la región opuesta, hasta encontrar una pareja de líneas que correspondan a los bordes de un mismo carril. Se considera que dos líneas $(\vartheta_{izda}, \rho_{izda})$ y $(\vartheta_{dcha}, \rho_{dcha})$ pertenecen al mismo carril cuando:

$$\begin{cases} |\vartheta_{izda} - \vartheta_{dcha}| < \varepsilon_\vartheta \\ |(|\rho_{izda} - \rho_{dcha}| - \Delta\rho_{ancho})| < \varepsilon_\rho \end{cases} \quad (4.4)$$

donde $\Delta\rho_{ancho}$ es el ancho del carril, y ε_ρ y ε_ϑ son las tolerancias admitidas en los parámetros ρ y ϑ , respectivamente. Sus valores se han fijado experimentalmente a: $\varepsilon_\rho = 1,5\text{m}$ y $\varepsilon_\vartheta = 4^\circ$. A cada pareja de líneas se le asigna como puntuación la suma de los votos de cada línea individual. En el caso de encontrar más de una pareja de líneas, se utiliza la de mayor puntuación.

En los siguientes fotogramas, las regiones se centran en los valores dados por la estimación del filtro de Kalman, y su tamaño se fija en seis veces la desviación típica de los parámetros ρ y θ (ver figura 4.22)³.

Este mecanismo es más rápido que la utilización de la Transformada Generalizada de Hough para buscar el modelo de carril recto (figura 4.20). Al tratarse de un modelo de tres parámetros, se necesita un acumulador de tres dimensiones que requiere un tiempo de cálculo mucho mayor. Por otro lado, el método de emparejamiento de líneas facilita la definición de tolerancias en el ancho y en la diferencia de orientaciones entre las dos líneas del carril.

Una vez que el carril ha sido detectado, se clasifican las dos líneas que lo componen en las tres clases ya descritas, y se buscan carriles adicionales de la misma manera que en el apartado anterior. Este proceso se muestra de forma gráfica en la figura 4.23.

4.4.4. Modificaciones en el mecanismo de votación de la Transformada de Hough

Pese a que el modelo de carretera recto puede considerarse válido en los 30–40 metros más próximos al vehículo, en curvas pronunciadas las líneas

³Según la desigualdad de Chebyshev, para cualquier distribución la probabilidad de que una variable aleatoria tenga un valor comprendido en un intervalo de seis veces la desviación típica, centrado en la media, es del 89%.

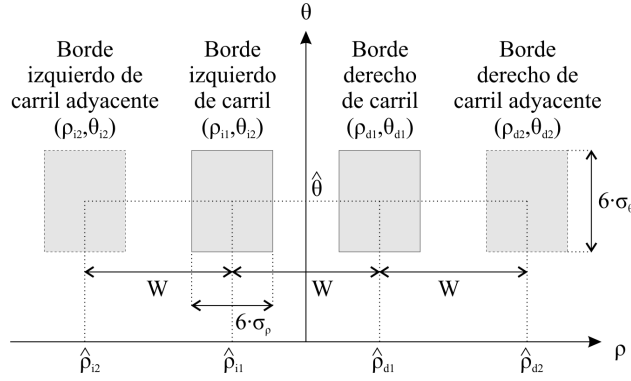


Figura 4.22: RDIs delimitadas por Kalman

detectadas por la transformada de Hough se ajustan como tangentes a las líneas viales, y dependiendo de la visibilidad y el estado de las mismas, el modelo de carril puede tender a ajustarse mejor en una zona próxima o alejada del vehículo. En el caso de que las líneas se ajusten en la zona más alejada, la máxima desviación de las líneas detectadas respecto a las marcas viales se dará en la zona más próxima, y el vehículo puede llegar a aparecer fuera del carril. En la figura 4.24 se ha exagerado este efecto para poder apreciarlo mejor.

Sin embargo, en la imagen capturada por las cámaras, el número de píxeles de marca vial que aparecen en la región próxima al vehículo es mucho mayor debido al efecto de la perspectiva. Por lo tanto, al generar la imagen en perspectiva inversa resulta más fiable, y además aparece más nítida, la región inferior (próxima al vehículo), que la superior. Por otro lado, a la hora de ajustar el modelo a la carretera, donde más precisión interesa es en la región más cercana al vehículo, pues ésta es la que el vehículo va a recorrer de forma inmediata.

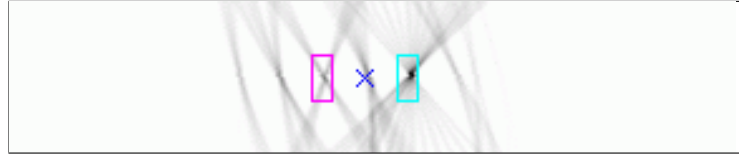
Por este motivo, se ha modificado el procedimiento de votación en la transformada de Hough, de forma que los votos que emiten los píxeles van perdiendo relevancia a medida que nos alejamos del vehículo.

Como se aprecia en la figura 4.25a, una distancia Δy de carretera se proyecta sobre el sensor de la cámara con un tamaño Δv . La relación $\Delta v/\Delta y$ va decreciendo a medida que la variable y se aleja del origen de coordenadas (es decir, del vehículo).

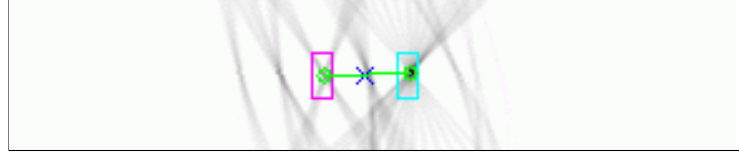
Para ponderar el valor del voto que emite un píxel, se define un peso w_i cuyo valor es proporcional al cociente $\Delta v/\Delta y$. De la figura 4.25b se puede deducir el siguiente sistema de ecuaciones:

$$\begin{cases} \frac{v}{f} = \tan \Delta\vartheta \\ \frac{H}{y} = \tan (\vartheta - \Delta\vartheta) \end{cases} \quad (4.5)$$

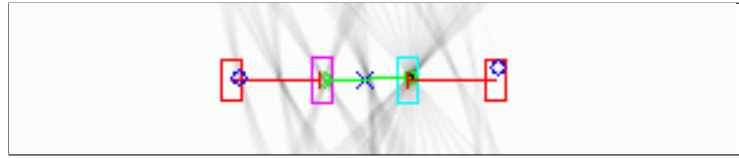
donde f es la distancia focal, ϑ es la inclinación de la cámara respecto de la



(a) RDIs para los bordes izquierdo y derecho



(b) Detección de un carril en las RDIs



(c) Nuevas RDIs para la detección de carriles adyacentes

Figura 4.23: Proceso de detección de carriles, mostrado sobre el acumulador de Hough.

horizontal, y H es la altura de la cámara.

Igualando $\Delta\vartheta$ en (4.5), y despejando la variable v , se obtiene (4.6):

$$v(y) = f \tan \left(\vartheta - \arg \tan \frac{H}{y} \right) \quad (4.6)$$

Finalmente, los pesos se definen como:

$$w(y_i) = k \frac{\Delta v}{\Delta y} = k \frac{v(y_{i+1}) - v(y_i)}{y_{i+1} - y_i} \quad (4.7)$$

donde k es una constante de proporcionalidad que permite controlar las cotas superior e inferior de los pesos.

Los experimentos han demostrado que la utilización de los pesos w_i han dado lugar a un significativo aumento de la eficacia, ya que se consigue un mejor ajuste del carril en la parte inferior de la imagen, reduciendo así las oscilaciones en los parámetros del carril estimado. La figura 4.26 muestra la curva de pesos utilizados en los experimentos.

4.4.5. Resultados

La figura 4.27 muestra algunos resultados obtenidos con este algoritmo, donde se aprecia que funciona razonablemente bien en tramos rectos. El tiempo de ejecución también es aceptable, siendo de unos 80–90 ms por fotograma en un Pentium IV, incluyendo la captura y rectificación de las imágenes es-

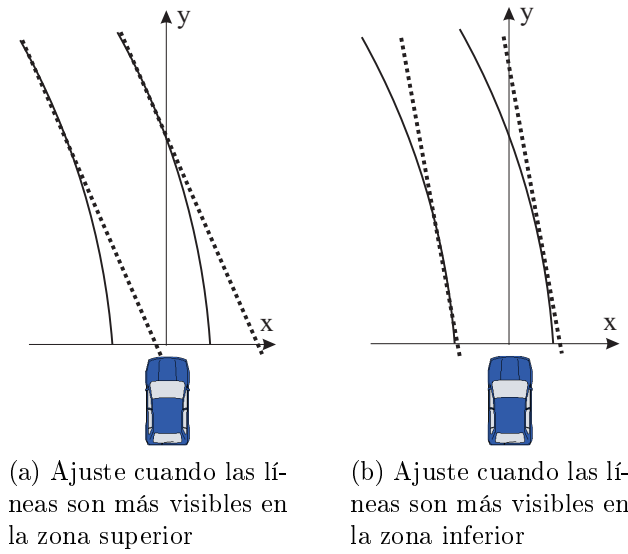


Figura 4.24: Problemas con el ajuste del modelo de carril en las curvas.

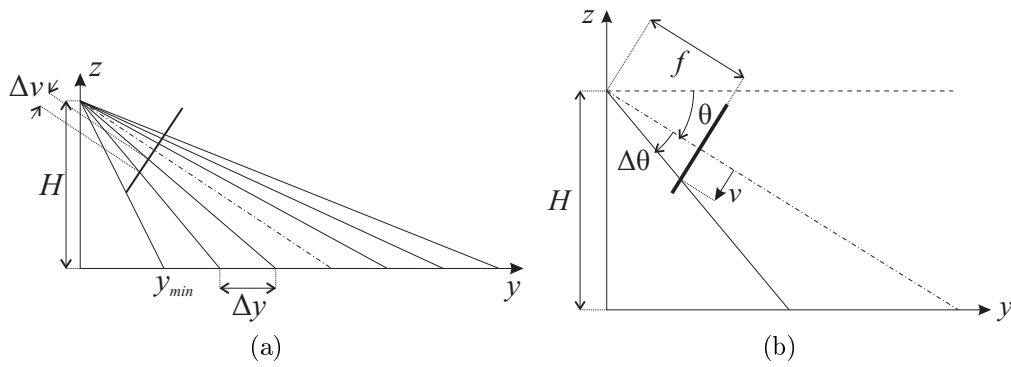


Figura 4.25: Cálculo de los pesos w_i .

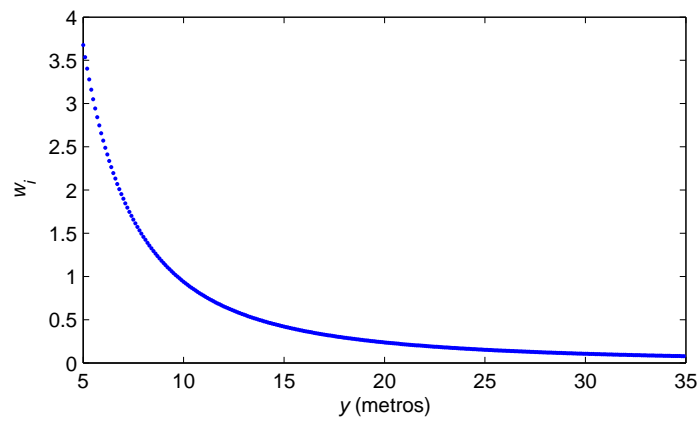


Figura 4.26: Pesos para $\Delta y = 0,1$ m/píxel, $\vartheta = 0,05$ rad, y $H = 1,18$ m.

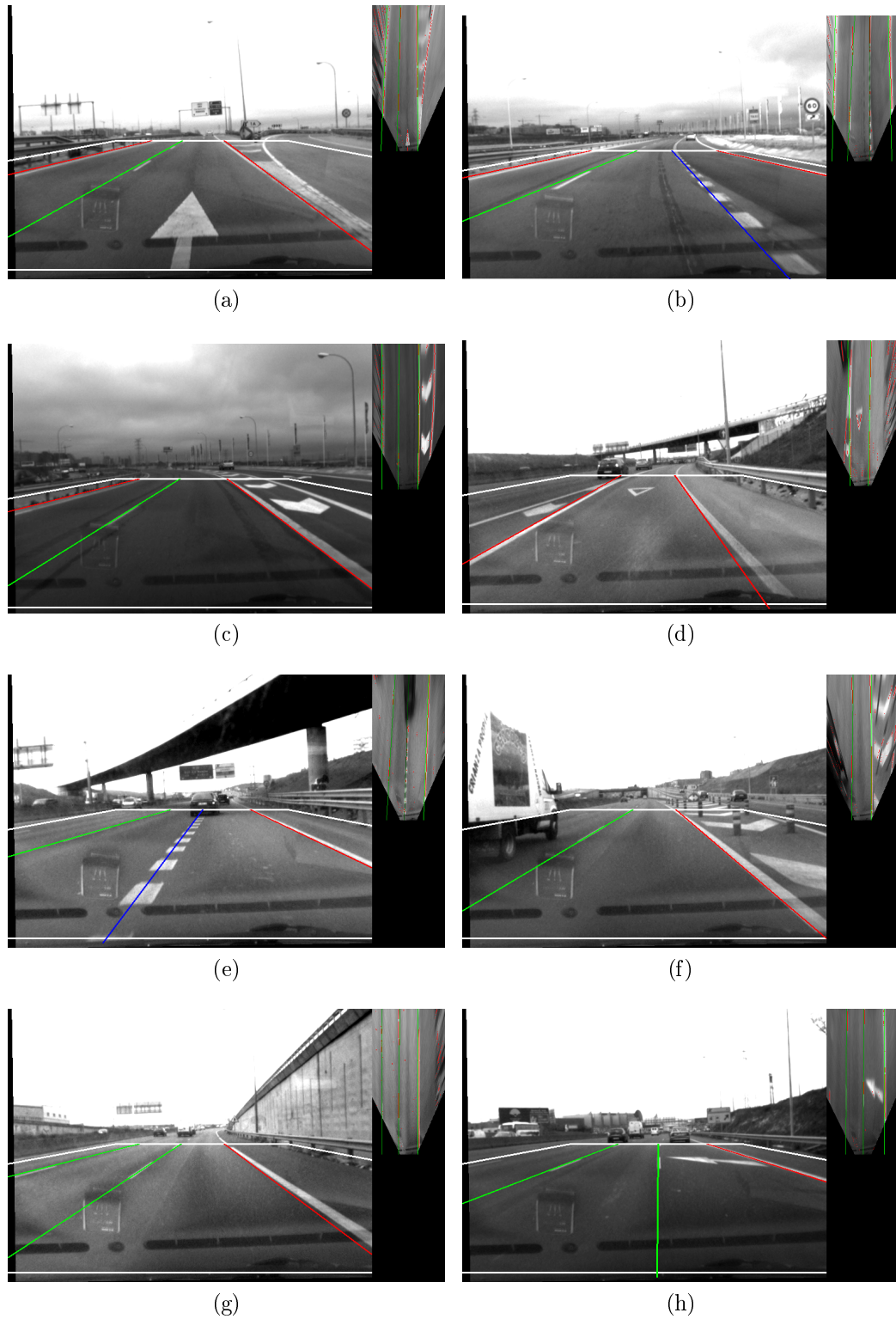


Figura 4.27: Ejemplos de detección de carriles. En cada ejemplo, la imagen en perspectiva inversa aparece junto a la esquina superior derecha. En la imagen original la zona remuestreada aparece delimitada por una línea blanca.

téreo, la Transformada de Hough y la clasificación de las líneas viales. Esto da una velocidad de procesamiento de 11–12 fps.

Si embargo, el modelo tiene dificultades para seguir el carril en los tramos curvos. La figura 4.27 muestra la evolución de los tres parámetros del modelo de carril a lo largo de una secuencia. Se aprecia, hacia el fotograma 930, una curva que se reconoce correctamente, pero en torno a los fotogramas 1080 y 1200 los parámetros representados presentan oscilaciones que corresponden a pérdidas del carril en una curva y en un cambio de carril, respectivamente.

Esto se debe a la simplicidad del modelo utilizado. Por una parte, las variables no son totalmente independientes, y por otra, el vehículo no poseía en ese momento ningún sensor de propiocepción, como acelerómetros o GPS, de manera que no era posible introducir variables de entrada en las ecuaciones de estado.

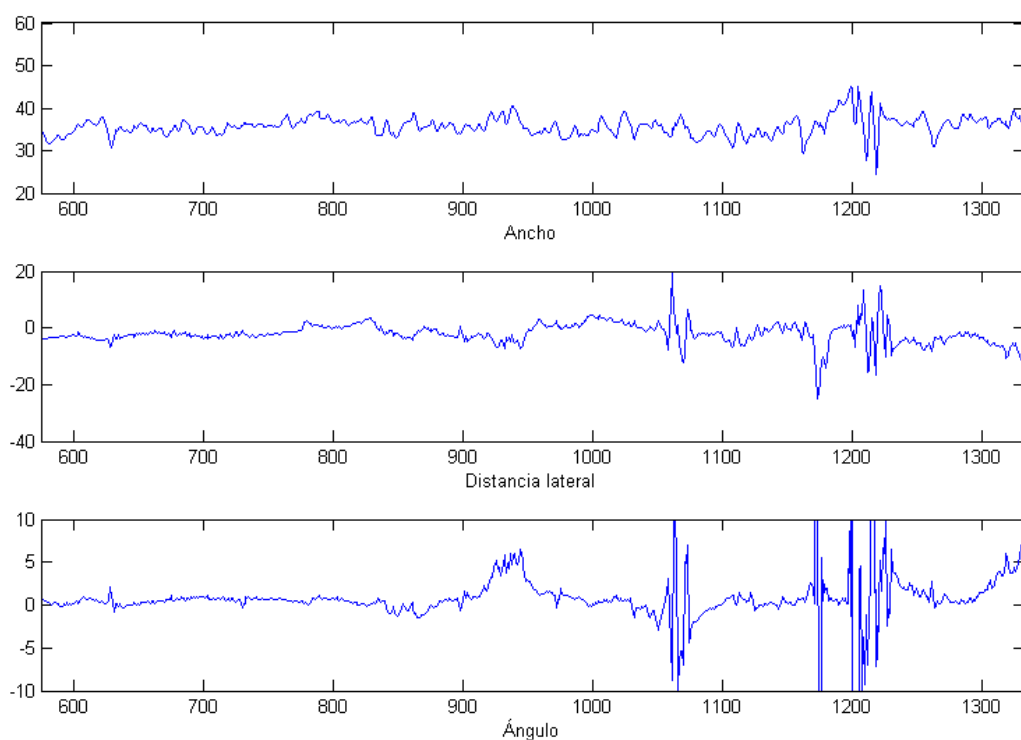


Figura 4.28: Detección de carriles con Kalman a lo largo de una secuencia

Por lo tanto, era necesario rediseñar el modelo de estados. Pero, teniendo en cuenta que el modelo recto no es sino un paso intermedio para el diseño de un modelo curvo, además de rediseñarlo, era deseable ampliarlo para incluir la curvatura del carril. Esto requería, además, un replanteamiento de las técnicas utilizadas en el procesamiento, ya que de seguir empleando las mismas, habría que implementar la Transformada de Hough para círculos que al necesitar un acumulador de tres dimensiones es considerablemente más lenta.

Michael Isard y Andrew Blake propusieron en 1998 un algoritmo para seguir objetos en entornos ruidosos (Isard, 1998). El algoritmo, denominado

CONDENSATION (*Conditional Density Propagation*), está basado en el filtro de partículas, del cual el filtro de Kalman es un caso particular. El seguimiento de formas se hace mediante un muestreo estadístico, de modo que ya no es necesaria la Transformada de Hough, y, a diferencia del filtro de Kalman, admite cualquier tipo de distribución estadística. Además, junto al algoritmo los autores proponen un mecanismo para identificar el modelo de estados de forma automática, que hace innecesaria tanto su deducción analítica como el estudio de la incertidumbre.

Dado que la propuesta de Isard y Blake resolvía, al menos en la teoría, los problemas antes reseñados, se optó por sustituir el filtro de Kalman por una implementación del filtro de partículas adaptada a este caso, incluir un procedimiento de identificación de sistemas, y redefinir el algoritmo de procesamiento para adecuarlo al nuevo algoritmo de seguimiento. La implementación de todas estas propuestas se describe en el próximo apartado.

4.5. Modelo de carretera con curvatura

La Norma 3.1-IC (BOE, 2000) sobre el trazado de carreteras contempla dos tipos de curvas. Las curvas circulares, que siguen la curva de la circunferencia, y las curvas de transición, que unen las anteriores con los tramos rectos, y que siguen la curva de la clotoide. En este trabajo se han aproximado estas curvas con una parábola, que tiene una expresión en coordenadas cartesianas más sencilla. La aproximación de una curva circular con una parábola se deduce en el apéndice C.

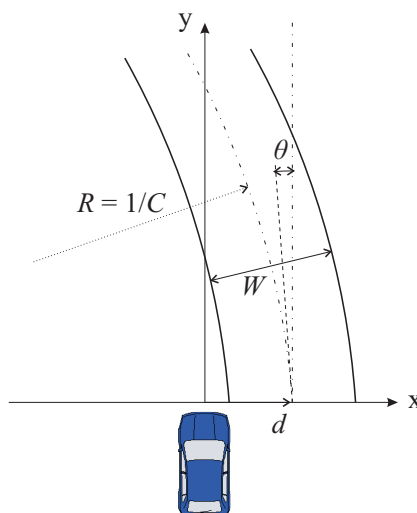


Figura 4.29: Modelo de carril con curvatura

El modelo propuesto de carril con curvatura está representado en la figura 4.29, y se compone de los siguientes cuatro parámetros:

- C : Curvatura, que se define como la inversa del radio de la curva ($C = 1/R$),
- θ : Orientación del vehículo respecto al eje del carril,
- d : Distancia al eje del carril,
- W : Ancho de carril.

Las líneas viales que delimitan el carril siguen el perfil parabólico definido por la ecuación 4.8, y se consideran como desplazamientos horizontales del eje del carril:

$$x(y) = \frac{C}{2} y^2 - \theta y - d - \frac{kW}{2} \quad (4.8)$$

donde k es un índice que referencia cada línea vial, y que sólo toma valores impares, positivos y negativos. El algoritmo de seguimiento considera hasta tres carriles, es decir, cuatro líneas viales que se representan por los valores $k = \{3, 1, -1, -3\}$. El valor $k = 0$ representaría el eje central del carril.

4.5.1. Seguimiento de carriles con un filtro de partículas

A diferencia del filtro de Kalman, los denominados *filtro de partículas* o *métodos de Monte Carlo secuenciales* no exigen distribuciones gaussianas ni modelos lineales, y permiten muestrear y seguir distribuciones multimodales con un número de muestras (denominadas partículas) reducido, ya que éstas tienden a concentrarse en torno a los máximos locales de la distribución de densidad de probabilidad. Además, los filtros de partículas tienen una mayor capacidad de recuperación ante pérdidas del objeto a seguir que el filtro de Kalman.

Para el seguimiento de carriles se ha escogido una técnica *SIR* (siglas en inglés de *Sampling Importance Resampling*) denominada algoritmo *ConDensation*, propuesta por M. Isard y A. Blake (Isard y Blake, 1998), y que ha sido diseñada para el seguimiento en tiempo real de formas complejas en secuencias de vídeo.

La dinámica del carril se ha modelado mediante un modelo autorregresivo de segundo orden (ecuación 4.9):

$$\mathbf{x}_t = A_2 \mathbf{x}_{t-2} + A_1 \mathbf{x}_{t-1} + D_0 + B_0 \mathbf{w}_t \quad (4.9)$$

donde \mathbf{x}_t es el vector de estado, compuesto por los cuatro parámetros del modelo de carril (figura 4.29), y \mathbf{w}_t es un vector de ruido gaussiano.

Estimación de la Función de Densidad de Probabilidad

El algoritmo ConDensation requiere una función que evalúe la bondad del ajuste de una hipótesis de carril (\mathbf{x}) a la observación (\mathbf{z} , que en este caso es la imagen analizada). Se trata por tanto de una expresión que aproxime la función de densidad de probabilidad de la observación $p(\mathbf{z}|\mathbf{x})$. En este trabajo la expresión utilizada se compone de dos términos:

- El primer término, F_1 , es una suma ponderada del número de píxeles de marcas viales que coinciden con los píxeles del carril (ecuación 4.10):

$$F_1 = \sum_{i=0}^N w_i \cdot I_{MV}(x_i, y_i) \quad (4.10)$$

donde:

- I_{MV} es la imagen de marcas viales (figura 4.30b), de tamaño $M \times N$,
- y_i y $x_i = x(y_i)$ son cada uno de los píxeles del carril, que se recorren con el índice i , y se calculan con la ecuación 4.8. $I_{MV}(x_i, y_i)$ es la intensidad de gris del píxel (x_i, y_i) .
- $w_i = w(y_i)$ son los pesos descritos en el apartado 4.4.4 que representa la relación entre el tamaño de un píxel en coordenadas del mundo y en coordenadas de la imagen. Su valor viene dado por la ecuación 4.7, y está representado en la figura 4.26.

Las coordenadas y_i y x_i representan el conjunto de píxeles que corresponden a la hipótesis de carril, \mathbf{x}_t , dibujada sobre la imagen en perspectiva inversa. Estas coordenadas están expresadas en el Sistema del Mundo, y siguen las ecuaciones 4.11 y 4.8, respectivamente, donde Δy es la altura de un píxel, y y_{min} es el valor de y correspondiente al píxel inferior de la imagen (figura 4.25a).

$$y_i = y_{min} + i \Delta y \quad (4.11)$$

- El segundo término, F_2 , mide la proximidad entre el carril y las marcas viales detectadas (ecuación 4.12):

$$F_2 = \sum_{i=0}^N w_i \cdot I_{DMV}(x_i, y_i) \quad (4.12)$$

donde:

- y_i , x_i y w_i son los mismos que para el término F_1 ,

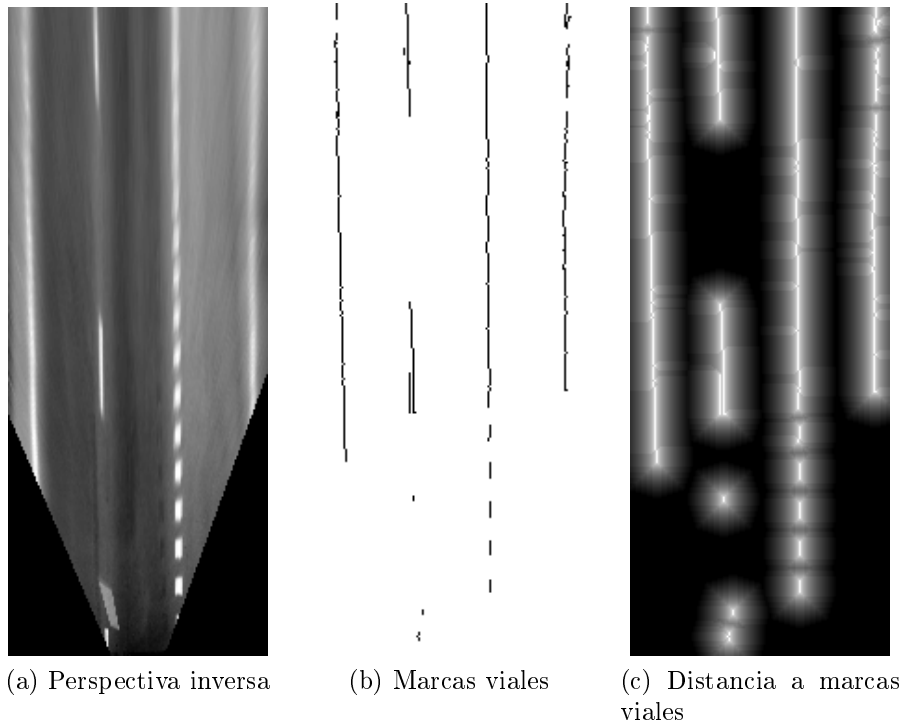


Figura 4.30: Imágenes utilizadas en el seguimiento del carril con el filtro de partículas.

- y I_{DMV} es la imagen de Distancia a Marcas Viales (figura 4.30c), que se obtiene aplicando la Transformada Distancia con caída exponencial a la imagen de marcas viales I_{MV} (figura 4.30b).

La función de densidad de probabilidad a posteriori se estima mediante una suma ponderada de los dos términos anteriores (ecuación 4.13):

$$F = k_1 \cdot F_1 + k_2 \cdot F_2 \quad (4.13)$$

donde k_1 y k_2 son constantes utilizadas para dar la misma importancia relativa a ambos términos.

La salida del algoritmo de seguimiento es la partícula \mathbf{x}_t que corresponde al máximo valor de F . En otras palabras, se toma como estimación la moda de la distribución de probabilidad. Esto provoca un cierto traqueteo en el seguimiento, debido a saltos bruscos de una partícula a otra, pero da resultados mucho más precisos que la media.

Aprendizaje del Modelo Dinámico

M. Isard, en su tesis doctoral (Isard, 1998) propone un mecanismo de aprendizaje automático de los parámetros del modelo dinámico del sistema, que consiste en utilizar primero un modelo sencillo que se introduce de forma

manual, para seguir una secuencia sencilla. A continuación, el modelo dinámico se ajusta por regresión a las observaciones obtenidas en esa secuencia. Este nuevo modelo se utiliza para seguir una secuencia más compleja, y se repite la regresión para obtener un modelo más refinado. Este proceso se realiza varias veces hasta obtener un modelo capaz de seguir adecuadamente los objetos en condiciones reales.

En esta tesis se ha aplicado esta técnica para el aprendizaje de los parámetros del modelo autorregresivo de segundo orden (ecuación 4.9): A_2 , A_1 , D_0 y B_0 . Primero se utilizó un modelo sencillo, en el que todas las variables eran independientes, para seguir el carril en una secuencia de un tramo recto seguido de un giro a la izquierda y luego otro a la derecha, sin tráfico ni cambios de carril. Con las observaciones de esta secuencia se minimizaron los errores del modelo dinámico, y con este nuevo modelo se siguió un carril en una secuencia que sí incluía tráfico y cambios de carril. Mientras que el primer modelo requirió 10.000 partículas para asegurar la convergencia, el último modelo aprendido tan sólo requiere 1.000.

4.5.2. Resultados

El seguimiento de carriles curvos mejora sensiblemente los resultados obtenidos con el seguimiento de carriles rectos, sin pérdida de velocidad. Todo el procesamiento se ejecuta a una velocidad de 12 fps en un procesador Pentium IV a 2.2GHz. El preprocesamiento, seguimiento y clasificación de líneas viales se realiza en aproximadamente 30 ms. El resto del tiempo se invierte en la captura y rectificación de las imágenes.

La figura 4.32 muestra las detecciones del algoritmo a lo largo de una secuencia de 2.500 fotogramas que incluye giros a la izquierda y a la derecha, dos rotondas, y dos cambios de carril, correspondiente a una carretera de enlace entre un municipio y una autovía. Aunque el algoritmo no ha sido diseñado para entornos urbanos, esta secuencia presenta mayor ruido que una autovía bien conservada, por lo que se considera una buena zona de pruebas. En el gráfico figura la evolución de los cuatro parámetros que definen el carril y la puntuación obtenida (medida tal como se explicó en el apartado 4.5.1).

La gran dispersión del parámetro angular se debe a las vibraciones de la cámara. El sistema estéreo está fijado al vehículo mediante un brazo flexible unido a una ventosa adherida al parabrisas, como se aprecia en la figura 4.31. La forma en que el brazo flexible está colocado da lugar a la transmisión de vibraciones a las cámaras, especialmente en sentido horizontal.

El principal caso de fallo en esta secuencia se da en las glorietas, que corresponden a las dos zonas sombreadas de la figura 4.32. Este algoritmo está pensado para funcionar en la red principal de carreteras, por lo que en las glorietas la curvatura de las líneas viales excede a los límites impuestos a los parámetros de carril, de modo que la estimación no puede ajustarse correctamente a las observaciones. Sin embargo, en estos casos, la puntuación obtenida por la estimación de carril se reduce significativamente, y se utiliza



Figura 4.31: Sistema de fijación de las cámaras al parabrisas

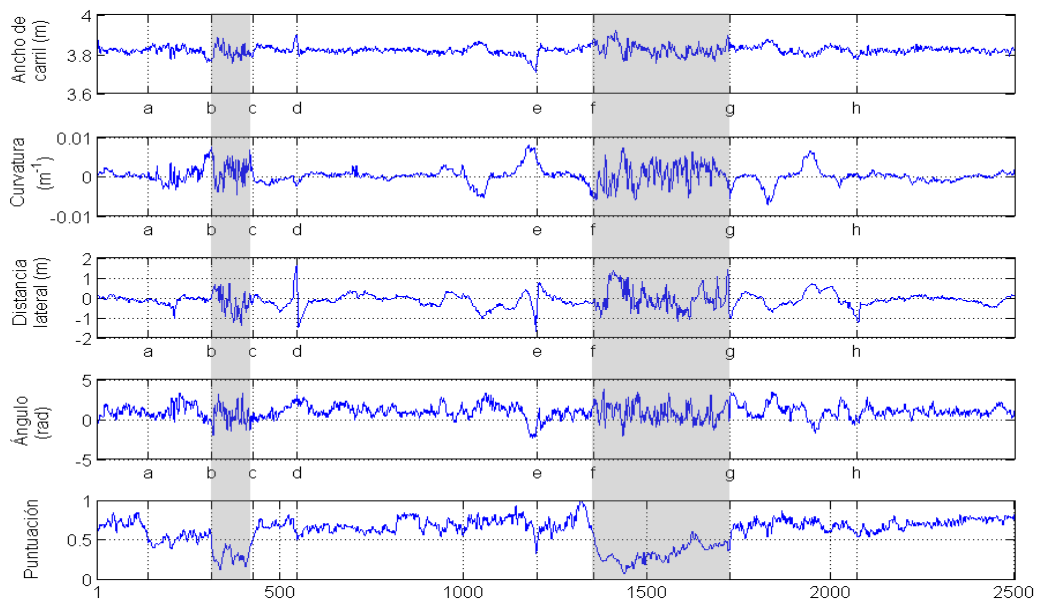


Figura 4.32: Secuencia correspondiente a la carretera de enlace de Leganés con la salida 28 de la autovía M-40, en Madrid.

como indicador de la fiabilidad del seguimiento. En consecuencia, el algoritmo suprime todas las alarmas cuando la puntuación está por debajo de 0,4.

El tramo entre los puntos (a) y (b) de la figura 4.32 presenta una gran dispersión en el parámetro de curvatura debido a la incorporación de un autobús justo enfrente del vehículo, por tanto ocluyendo las marcas viales casi por completo.

Los puntos (d) y (e) son cambios de carril detectados correctamente. El tramo comprendido entre (d) y (e) contiene un giro a la izquierda seguido de otro a la derecha, coherente con lo que se observa en el gráfico de curvatura. El cambio de carril tuvo lugar durante un giro a la derecha.

El tramo entre (g) y (h) de nuevo contiene dos curvas. El punto (h) representa otro modo de fallo, que se corresponde con la figura 4.33f, cuando la estimación de carril siguió la línea de salida hasta que apareció en la imagen

la línea discontinua de incorporación. El algoritmo confundió este caso con una salida de carril y lanzó una falsa alarma.

La figura 4.33 muestra algunos ejemplos, en los que los puntos negros representan los carriles detectados. La clasificación de marcas viales se representa cambiando el grosor y espaciado entre puntos.

La figura 4.33a presenta una carretera de tres carriles, mientras que en las figuras 4.33b y 4.33e muestran carreteras de dos carriles. La figura 4.33h corresponde a una carretera de dos carriles, pero al ser las líneas viales continuas, el algoritmo no busca carriles adyacentes.

La figura 4.33g muestra una carretera de dos carriles en el que sólo uno de ellos es detectado. Este fallo se debe a la curvatura vertical de la carretera al dificultar ésta que el modelo se ajuste bien a las líneas izquierda y derecha simultáneamente. El modelo tiende a seguir la línea de la derecha que al ser continua contiene más píxeles de marca vial. La línea izquierda correspondiente a la estimación de carril se aleja de la posición de la verdadera, y el perfil de línea extraído para la clasificación ya no corresponde a la línea vial sino al asfalto. Como se verá más adelante, por defecto, las líneas no detectadas como discontinuas de alguna clase se clasifican como continuas.

La figura 4.33e es un ejemplo de las ventajas de este algoritmo. Aunque un vehículo oculte casi completamente el carril izquierdo, los dos carriles existentes son detectados gracias a que la línea vial de separación es clasificada como discontinua.

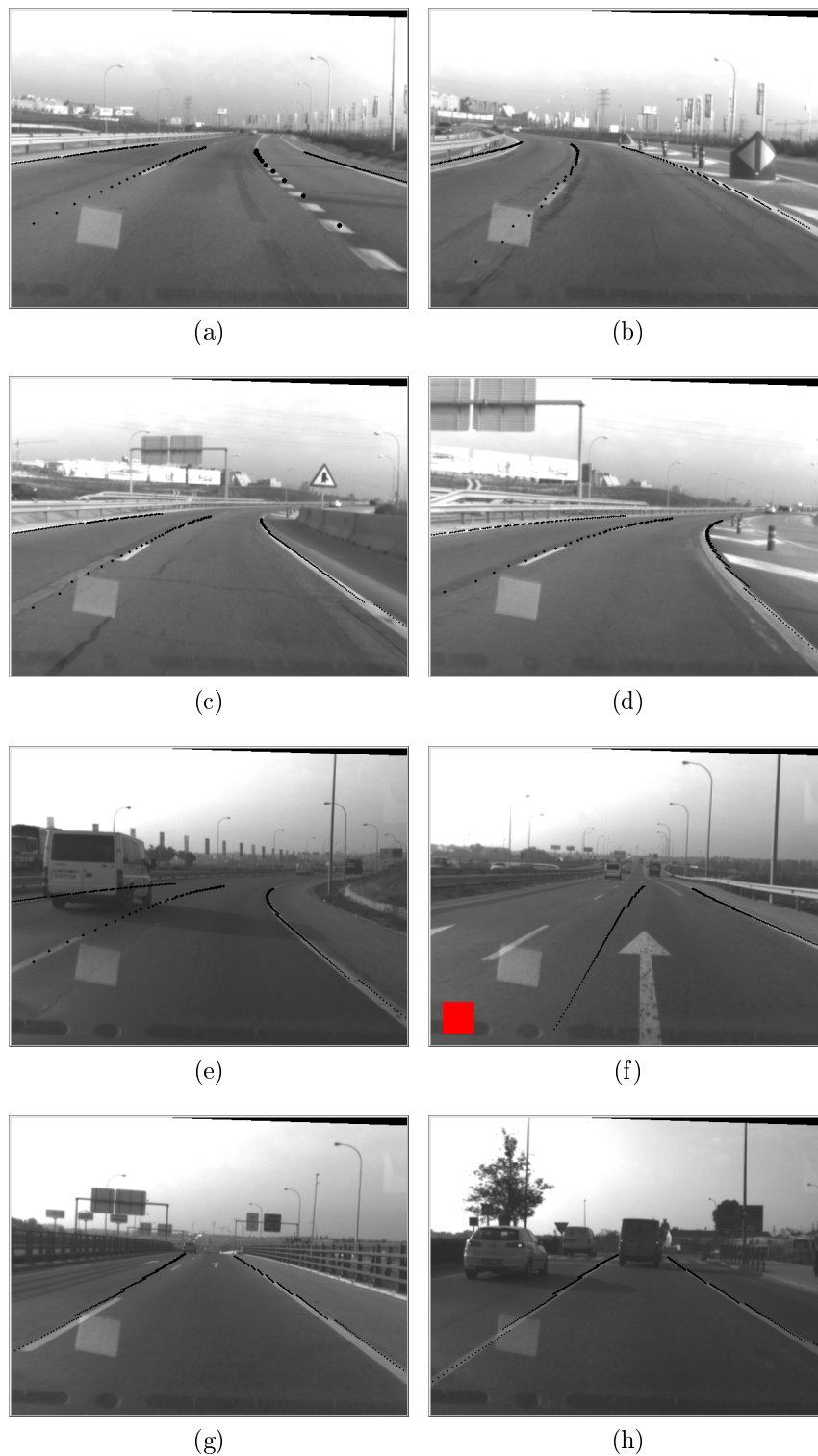


Figura 4.33: Algunos ejemplos de la secuencia ilustrada en la figura 4.32. Los puntos finos representan líneas continuas, los gruesos y de mayor espaciado representan líneas de entrada/salida, y los de tamaño y espaciado medios corresponden a líneas discontinuas (ver capítulo 5).

Capítulo 5

Clasificación de líneas viales longitudinales y Sistema de Alerta por Salida de Carril

5.1. Clasificación de líneas viales longitudinales

5.1.1. Introducción

La clasificación de líneas viales aporta una información del entorno que permite:

- *Detectar de forma automática el número de carriles de la carretera.* Por ejemplo, la presencia de una línea discontinua a la derecha revela la existencia de un carril en ese mismo lado.
- *Reconocer el tipo de vía por el que se circula.* Conocer el tipo y número de carriles en ocasiones puede permitir distinguir entre una carretera convencional de doble sentido y una autovía, por ejemplo.
- *Detectar maniobras incorrectas.* Conocer el tipo de líneas viales permite suprimir las alertas cuando se señala con el intermitente una maniobra permitida, y mantenerlas si esa maniobra es incorrecta, como por ejemplo el cruce de una línea continua, o un cambio de carril sin señalar.

El reconocimiento del tipo de líneas viales de separación de carriles aporta una información que facilita la evaluación de la peligrosidad de las maniobras, permitiendo, por ejemplo, distinguir entre el cruce de una línea continua y una discontinua. A pesar de esto, el tratamiento que ha recibido este tema en la literatura es escaso, a juicio del autor, por dos razones: primero, que hasta ahora en el ámbito de la asistencia a la conducción había otros temas más urgentes que resolver, y segundo, que desafortunadamente en la actualidad no existe una legislación internacional que regule la presencia y forma de las

líneas viales. Esto probablemente ha desalentado el desarrollo de este tipo de sistemas, ya que su aplicación queda reducida al ámbito nacional.

En España, las marcas viales están reguladas por la norma 8.2-IC “Marcas viales” (BOE, 1987). En la página web www.carreteros.org puede encontrarse una versión HTML de la norma, más cómoda de consultar en línea (carreteros.org, 1987).

La tabla 5.1 resume todos los tipos de marcas viales longitudinales contemplados por la norma, que hacen un total de 21, a saber: 12 clases de líneas discontinuas, 6 de líneas continuas, y 3 de líneas discontinuas adosadas a continuas.

Tabla 5.1: Tipos de marcas viales longitudinales según la Norma 8.2-IC

Significado	Marca
Marcas longitudinales discontinuas:	
Para separación de carriles normales	M-1.1, M-1.2, M-1.3
Para separación de carriles reversibles	M-1.4, M-1.5
Para separación de carriles de entrada o de salida	M-1.6, M-1.7
Para preaviso de bifurcación	M-1.8
Para preaviso de marca continua	M-1.9, M-1.10
Para borde de calzada	M-1.11, M-1.12
Marcas longitudinales continuas:	
Para separación de carriles en el mismo sentido	M-2.1
Para separación de sentidos	M-2.2, M-2.3
Para separación de carril especial o de entrada o de salida	M-2.4
Para borde de calzada	M-2.5, M-2.6
Marcas longitudinales continuas adosadas a discontinuas:	
Para regulación de adelantamiento	M-3.1, M-3.2, M-3.3

Existen limitaciones físicas para el reconocimiento de todo este conjunto de líneas, fundamentalmente debidas al sistema de percepción empleado. El ruido de la imagen dificulta la extracción del perfil de la línea, y pequeños errores en la detección de marcas viales provocan discontinuidades que distorsionan el análisis de frecuencias. Éstos errores no tendrían mayor importancia si se analizase una longitud grande de la línea, pero esto no es posible salvo que se utilicen cámaras de alta resolución, ya que a resoluciones convencionales de 640×480 las líneas viales aparecen difuminadas en el entorno lejano, de modo que a partir de cierta distancia el perfil percibido por las cámaras deja de ser válido para un análisis de frecuencias.

Por otra parte, el tener un conjunto limitado de líneas permite elevar la robustez de la clasificación y le da al sistema mayor aplicabilidad en el extranjero.

Por lo tanto, de todo este conjunto de líneas, el algoritmo de clasificación propuesto es capaz de distinguir entre los siguientes tipos de línea:

- *Marcas longitudinales discontinuas para separación de carriles normales* (figura 5.1). A estos tipos de marcas se les asignará la clase **línea discontinua**, y corresponden a los siguientes tipos:

- M-1.1 (para vías con velocidad máxima superior a 100 km/h): franjas de 5 m espaciadas 12 m.
- M-1.2 (para vías con velocidad máxima entre 60 y 100 km/h): franjas de 3,5 m espaciadas 9 m.

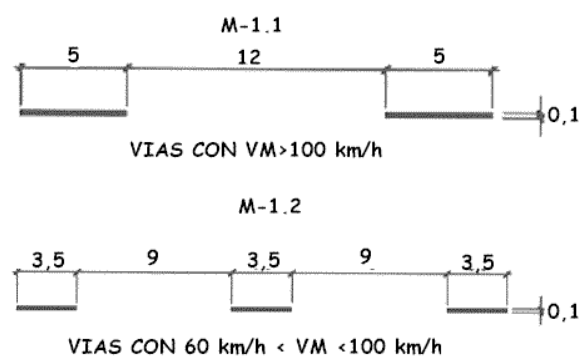


Figura 5.1: Líneas discontinuas.

- *Marcas longitudinales discontinuas para separación de carriles de entrada o de salida, o de preaviso de bifurcación*, que se reconocerán como de la clase **línea de entrada/salida** (figura 5.2). Todas ellas son franjas de 1 m espaciadas 1 m, y corresponden a los tipos que sólo varían en el ancho de la marca :

- M-1.6: para separación de carril de entrada o de salida en vía con velocidad máxima superior a 100 km/h.
- M-1.7: para separación de carril especial o carril de entrada o de salida en vía con velocidad máxima inferior o igual a 100 km/h.
- M-1.8: para preaviso de bifurcación.

- *Marcas longitudinales continuas*, que serán reconocidas como **línea continua** (figura 5.3). Corresponden a los tipos de marca:

- M-2.1: para separación de carriles en el mismo sentido.
- M-2.2: para separación de sentidos en calzadas de dos o tres carriles.
- M-2.3: para separación de sentidos en calzadas de cuatro o más carriles.
- M-2.4: para separación de carril especial o carril de entrada o salida.

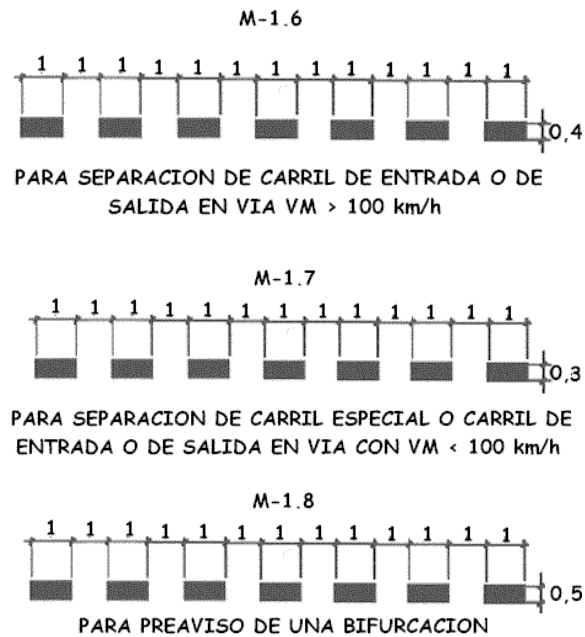


Figura 5.2: Líneas de entrada/salida.



Figura 5.3: Líneas continuas.

- M-2.5: para borde de calzada en vías con velocidad máxima superior a 100 km/h (borde izquierdo).
- M-2.6: para borde de calzada en vías con velocidad máxima inferior o igual a 100 km/h.

5.1.2. Análisis frecuencial de líneas viales

El algoritmo se basa en la realización de un análisis del perfil de la línea vial en el dominio de la frecuencia. Para ello, se aplica la transformada de Fourier, que para señales discretas está definida por la ecuación 5.1:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} k n}; \quad k = 0, \dots, N-1 \quad (5.1)$$

donde x es la variable discreta, N el número de muestras, y X su transformada de Fourier.

Como sólo interesa saber qué frecuencias hay presentes, se puede prescindir de la información de la fase, y trabajar con la potencia espectral $\|X_k\|^2$.

Las primeras decisiones que hay que tomar son la longitud del perfil de la línea que es necesario analizar, y la resolución con que se debe muestrear.

Según la norma, la línea discontinua de separación de carriles de mayor período es la línea M-1.1, que se compone de tramos de 5 metros separados una distancia de 12 metros, lo que da un período de 17 metros. Por lo tanto, será necesario extraer al menos 17 metros de perfil para asegurar que se muestrea al menos un periodo completo en todos los casos.

El análisis frecuencial se va a realizar sobre señales cuadradas. Aunque estas señales necesitan un número infinito de frecuencias para su reconstrucción completa, afortunadamente no es la reconstrucción de la onda el problema que nos ocupa, sino su clasificación en diversos tipos. En los siguientes apartados se realiza un estudio de viabilidad del algoritmo, partiendo de señales generadas artificialmente, con las que se comprueba la posibilidad de discriminar entre distintas marcas viales utilizando un número limitado de armónicos, a la vez que se analiza la influencia que tienen en el espectro de frecuencias la longitud de línea vial, su posición, y el ruido de la imagen. En el estudio se supondrá una frecuencia de muestreo correspondiente a la resolución de la imagen en perspectiva inversa, cuyo tamaño de píxel equivale a 10 cm de carretera.

Influencia de la distancia analizada

En este apartado se trata de estudiar qué longitud de línea vial debe analizarse para que el espectro de frecuencias no se vea sensiblemente afectado por la fase, entendida ésta como la distancia que hay entre el origen de coordenadas y el inicio del primer tramo pintado de la línea (figura 5.4).

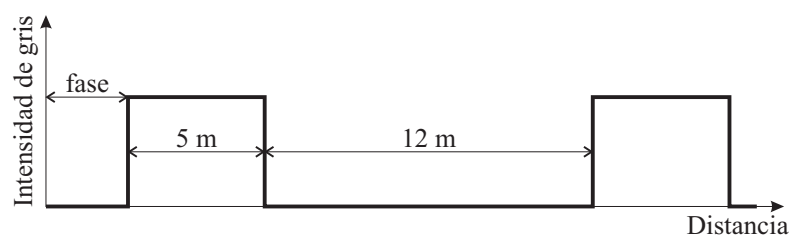


Figura 5.4: Fase de la línea vial M-1.1.

La figura 5.5, muestra los espectros de la marca M-1.1 (la de mayor periodo) para tres longitudes de perfil (20, 35 y 50 m) y dos fases diferentes. Inspeccionando las gráficas, vemos que para el caso de 20 metros el espectro de frecuencias varía ostensiblemente al modificar la fase. A partir de los 35 m, la distancia analizada supera el doble del periodo máximo, de modo que el espectro de frecuencias apenas presenta variaciones. El algoritmo de seguimiento de carriles (capítulo 4) explora una distancia de 35 m, que a tenor de estos resultados parece ser también una longitud válida para la clasificación de líneas viales.

Influencia del ruido

Se estudia ahora la sensibilidad del espectro de frecuencias al ruido, para los casos de las cuatro marcas viales más representativas (M-1.1 y M-1.3 para las líneas *discontinuas*, M-1.6 para las de *entrada/salida*, y M-2.1 para las *continuas*). En cada caso se altera la presencia de cada píxel con una cierta probabilidad, siendo cuatro los niveles de ruido analizados (0 %, 10 %, 20 % y 30 %).

Los resultados se muestran en las figuras 5.6 y 5.7, donde puede apreciarse que con un 30 % de ruido los picos de frecuencia característicos de cada línea empiezan a desaparecer. También se puede concluir de estos experimentos que teóricamente bastaría con analizar las 20 primeras frecuencias para determinar el tipo de marca vial de que se trata, aunque en la práctica será necesario ampliar el rango hasta la frecuencia 30, como se verá en el apartado 5.1.3.

Influencia del perfil de los datos

Otro aspecto importante del análisis frecuencias es la obtención de la señal que será objeto del análisis. Para esto existen dos alternativas:

1. Analizar el perfil de la línea en intensidad de gris (figura 5.9), extrayéndolo de la imagen remuestreada (imagen 5.8b).
2. Analizar el perfil de la línea en binario (figura 5.10), extraído de la imagen binaria de marcas viales detectadas (imagen 5.8c).

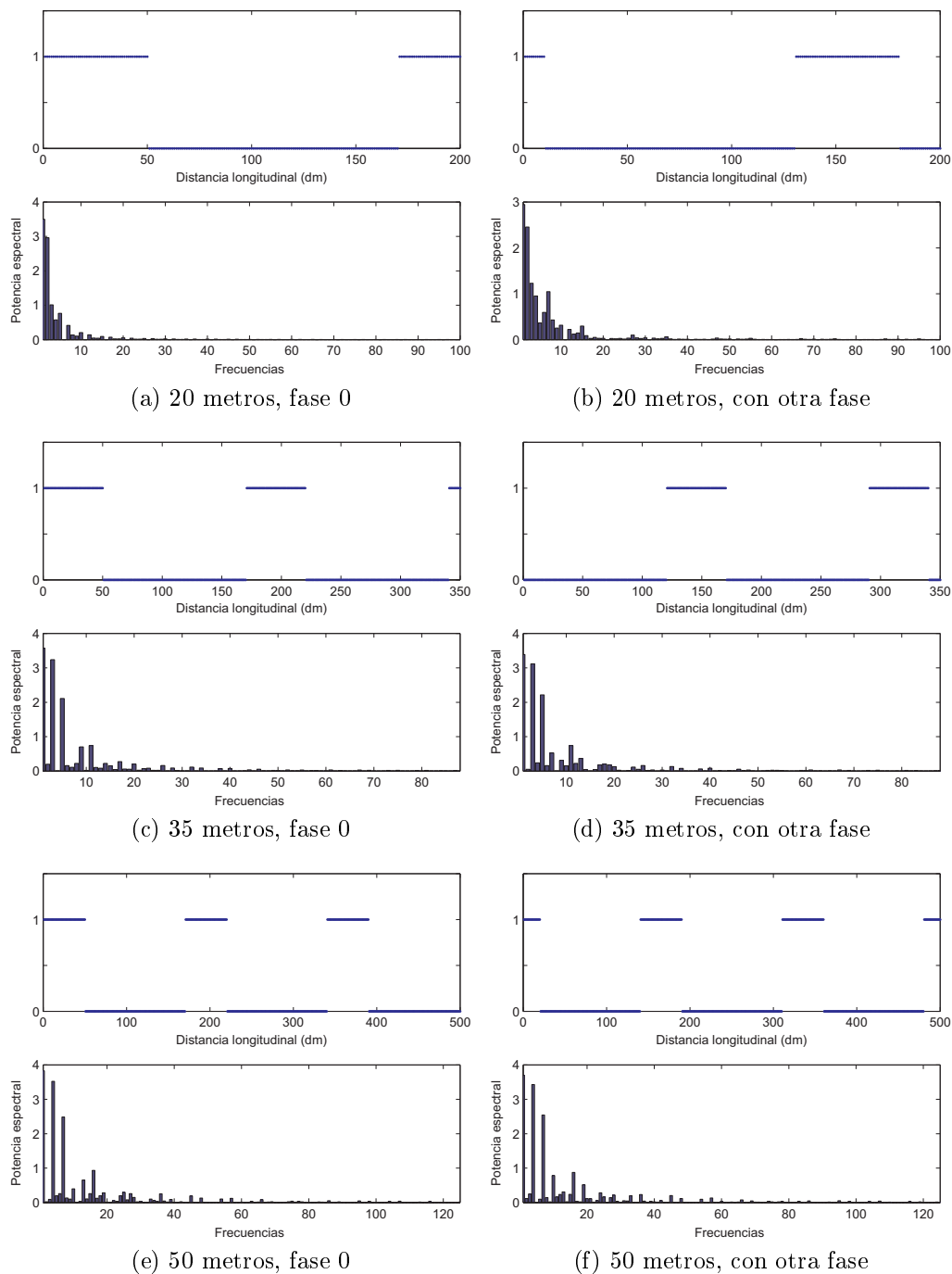


Figura 5.5: Influencia de la fase en el espectro de frecuencias, para la marca M-1.1.

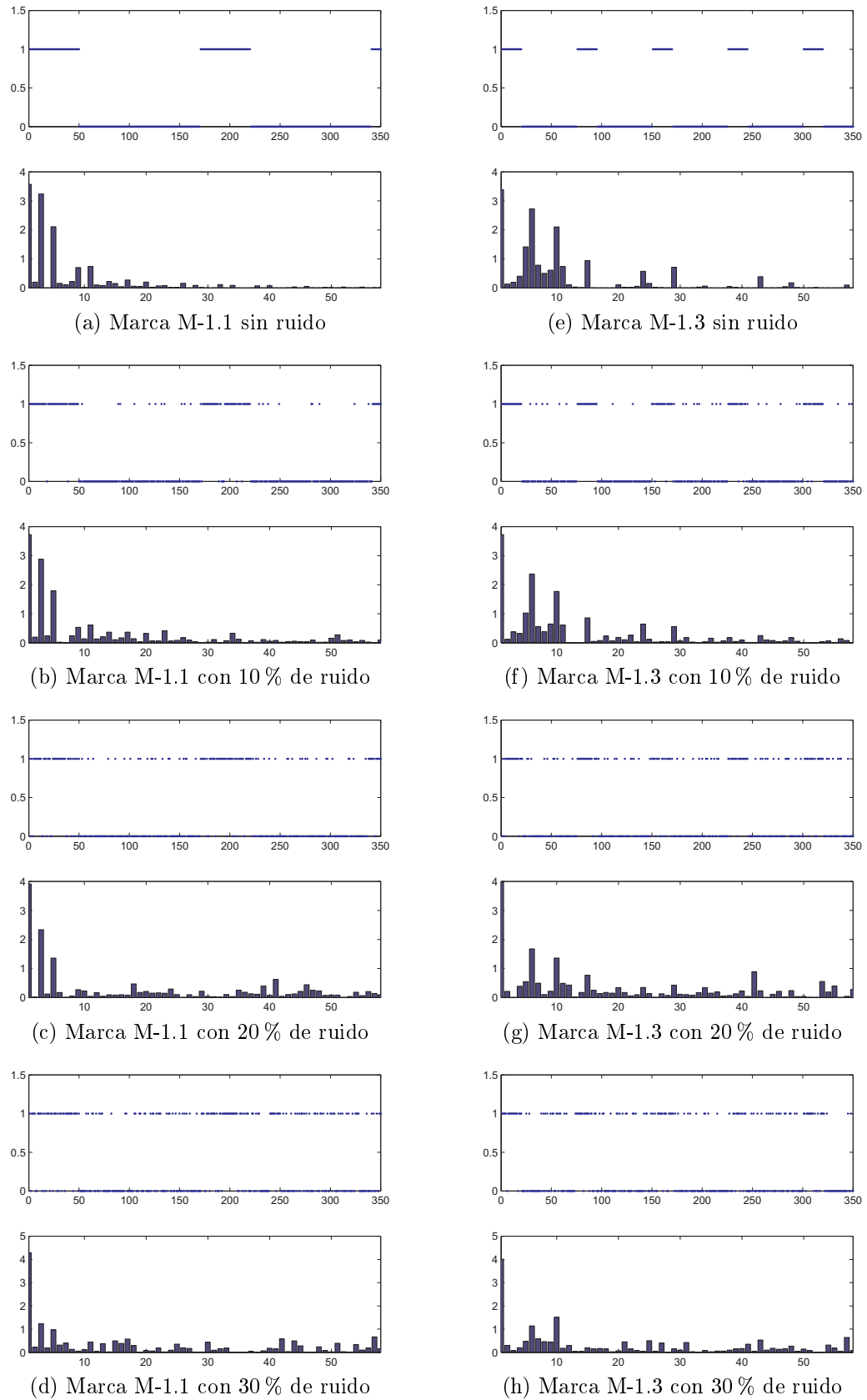
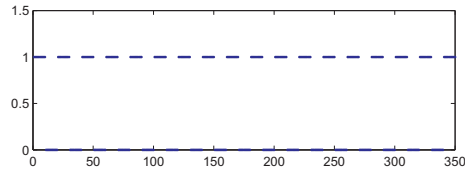
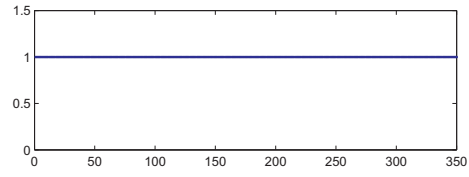


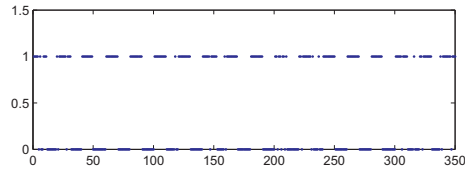
Figura 5.6: Influencia del ruido en el espectro de frecuencias (I).



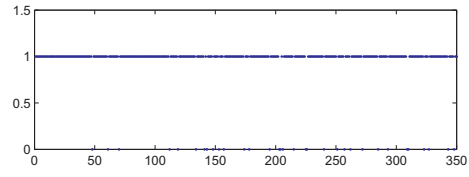
(a) Marca M-1.6 sin ruido



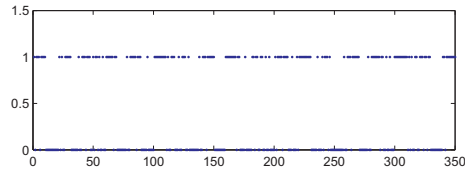
(e) Marca M-2.1 sin ruido



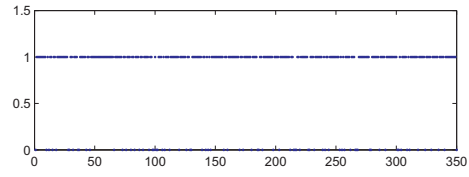
(b) Marca M-1.6 con 10 % de ruido



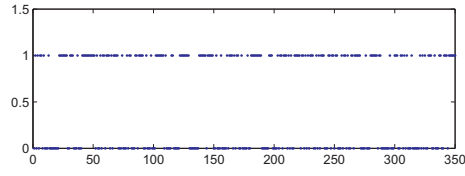
(f) Marca M-2.1 con 10 % de ruido



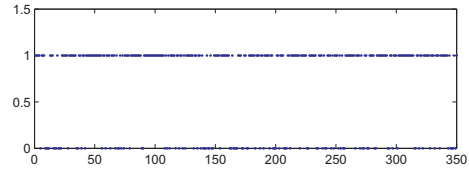
(c) Marca M-1.6 con 20 % de ruido



(g) Marca M-2.1 con 20 % de ruido

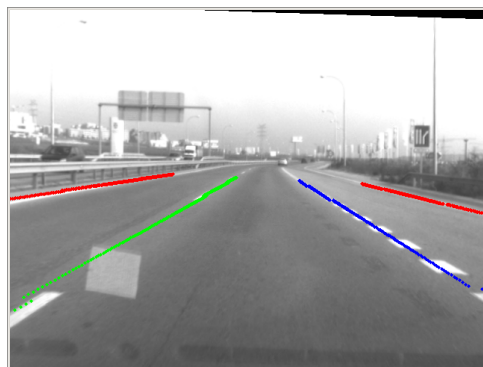


(d) Marca M-1.6 con 30 % de ruido

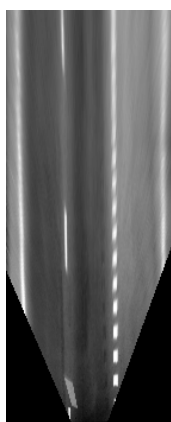


(h) Marca M-2.1 con 30 % de ruido

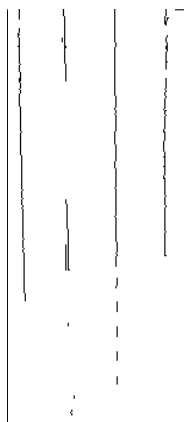
Figura 5.7: Influencia del ruido en el espectro de frecuencias (y II).



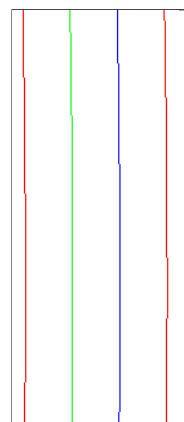
(a) imagen original con la detección y clasificación de líneas viales



(b) perspectiva inversa



(c) marcas viales detectadas



(d) líneas viales detectadas y clasificadas

Figura 5.8: Señales de partida para el análisis en frecuencias, y resultado de la detección y clasificación de líneas (en rojo las *continuas*, en verde las *discontinuas* y en azul las de *entrada/salida*).

El análisis del perfil de la línea en intensidad de gris tiene la ventaja de no verse afectado por el proceso de detección de marcas viales, con sus posibles errores, pero a cambio presenta una gran dependencia tanto de las condiciones ambientales como de las variaciones en el estado y la iluminación a lo largo de la línea vial. Por otro lado, necesita un ajuste muy preciso de la detección a la marca vial en toda su longitud, ya que de lo contrario el algoritmo extraería un perfil de intensidad correspondiente a regiones del asfalto no pintadas.

La figura 5.9, muestra los perfiles de las líneas en niveles de gris, y sus espectros de potencia en escala logarítmica. Las líneas (a)-(b)-(c)-(d) son las de la figura 5.8d, en sentido de izquierda a derecha, de las cuales (a) y (d) son *continuas*, (b) es *discontinua* y (c) es de *entrada/salida*. Aquí se observa con claridad la pérdida de resolución de la carretera en el entorno lejano al vehículo por efecto de la perspectiva. Esto puede considerarse como un ruido

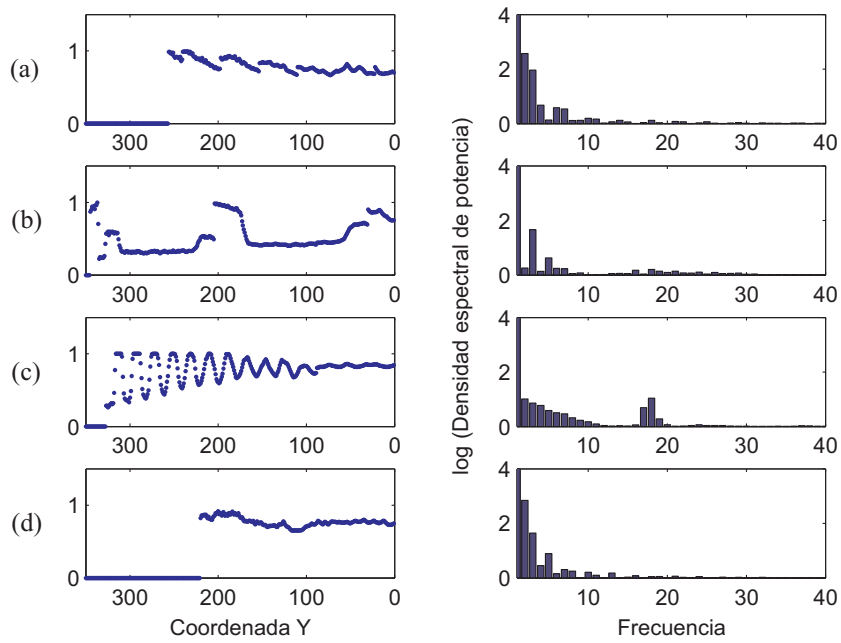


Figura 5.9: Espectro de potencia del perfil de intensidades en niveles de gris; (*izquierda*) perfiles de las líneas extraídos de la figura 5.8b, de izquierda a derecha; (*derecha*) densidad espectral de potencia

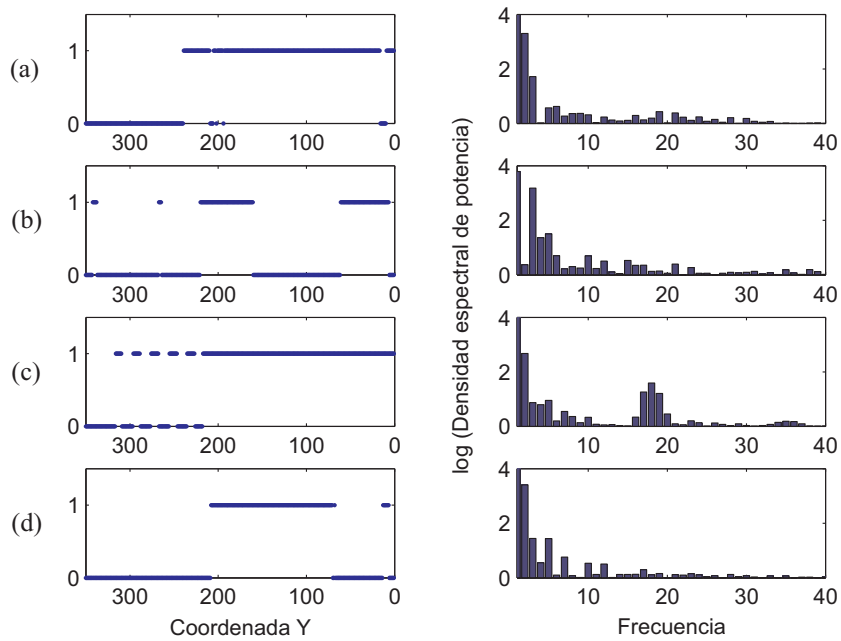


Figura 5.10: Espectro de potencia del perfil de intensidades en binario; (*izquierda*) perfiles de las líneas extraídos de la figura 5.8c, de izquierda a derecha; (*derecha*) densidad espectral de potencia

que afecta principalmente a las líneas de *entrada/salida*, que van difuminándose a medida que aumenta la distancia hasta presentar la apariencia de una línea continua.

El análisis del perfil de la línea en binario, aunque se ve afectado por el posible ruido introducido durante la detección de marcas viales, es más independiente de las condiciones de iluminación y menos sensible al estado de las marcas viales. Además permite establecer tolerancias en la extracción del perfil, de forma que es suficiente que un píxel detectado como marca vial esté próximo a la detección de la línea vial para considerarse como perteneciente a ella. De esta manera, el algoritmo de clasificación de líneas puede funcionar aún en el caso de que el ajuste del carril detectado no sea perfecto.

La figura 5.10 presenta el perfil de las líneas obtenido a partir de la detección de marcas viales (columna de la izquierda) y su densidad espectral de potencia (columna de la derecha). En este caso se aprecian con mayor claridad que en el anterior unos picos en los espectros de potencia de la línea *discontinua* (b) y de la línea de *entrada/salida* (c), situados en las frecuencias características de cada tipo de línea. No aparecen picos significativos para las líneas *continuas* (a) y (d).

5.1.3. Algoritmo de Clasificación

Una vez comprobado que el espectro de potencias de Fourier puede ser un buen discriminante para distinguir entre las tres clases de líneas propuestas, y que ésto puede hacerse utilizando información que se puede extraer fácilmente a partir del procesamiento realizado para la detección y seguimiento de carriles, es decir, sin necesidad de modificar ni la resolución de la imagen ni su tamaño (del que depende la longitud de línea analizada), el siguiente paso consiste en definir el algoritmo de clasificación para secuencias de vídeo reales, y comprobar su funcionamiento.

En el estudio anterior ha quedado confirmado que cada clase de línea produce unos picos en el espectro de Fourier a determinadas frecuencias. La deducción del rango de frecuencias característico de cada marca vial se ha realizado experimentalmente. Inspeccionando el espectro de Fourier para las líneas detectadas en una secuencia real de carretera se han podido definir los umbrales, tanto en frecuencia como en magnitud, que se relacionan en la tabla 5.2. En la figura 5.11 puede verse un ejemplo en el que aparecen las tres clases de líneas, y donde se señalan en color los umbrales en frecuencias y en magnitud.

Finalmente, el algoritmo de clasificación propuesto consiste en calcular, para cada línea detectada, las 30 primeras frecuencias del espectro de potencias, y seguir los siguientes pasos:

1. Si existe un pico en la densidad espectral de potencia con valor superior a 0.60 (en escala logarítmica), entre las frecuencias 20 y 29, se trata de una **línea de entrada/salida**.

Tabla 5.2: Frecuencias y umbrales en la clasificación de líneas viales

Línea	Frecuencia mínima	Frecuencia máxima	Umbral en el espectro de potencias (escala logarítmica)
Discontinua	3	5	1.50
Entrada/Salida	20	29	0.60
Continua	0	0	4.5

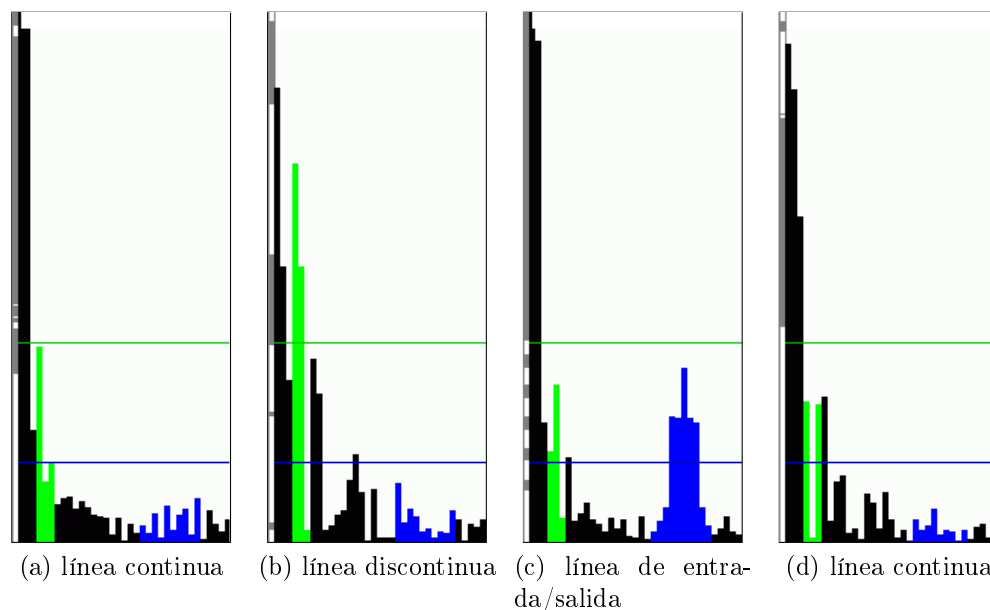


Figura 5.11: Espectro de potencia para las líneas viales de la figura 5.8. Las frecuencias características, así como los umbrales en magnitud (las líneas horizontales) se distinguen por el color: verde para las *discontinuas*, y azul para las de *entrada/salida*.

2. Si el valor de la densidad espectral de potencia para la frecuencia 0 es superior a 4.5, se trata de una **línea continua**.
3. Si existe un máximo local entre las frecuencias 3 y 5, con valor superior a 1.5, se trata de una **línea discontinua**.
4. Si no se cumplen las condiciones anteriores, se asume que se trata de una **línea continua** en la que las marcas viales no se han detectado correctamente debido al ruido o a defectos en la pintura.

El orden de estos pasos no es arbitrario. El hecho de comprobar primero la condición de línea continua, antes de discontinua, se debe a que ocasionalmente los defectos de pintura o la suciedad de la vía introducen en la señal frecuencias en el rango de las líneas discontinuas. Se ha observado que en estos casos se conserva un valor alto de continua (frecuencia 0), que permite diferenciar ambas situaciones.

5.1.4. Resultados

Las figuras 5.12, 5.13, 5.14 y 5.15 muestran ejemplos extraídos de secuencias reales. Todas ellas muestran la imagen obtenida por la cámara con las marcas viales detectadas superpuestas (a), la imagen en perspectiva inversa (b), y la densidad espectral de potencia (c). Las líneas clasificadas como *continuas* aparecen en rojo, las *discontinuas* en verde, y las de *entrada/salida* en azul.

En las figuras, el análisis de Fourier muestra el espectro de hasta cuatro marcas viales (dos a la izquierda y dos a la derecha del vehículo). De nuevo, las frecuencias de color verde son las características de las *líneas discontinuas* mientras que las de color azul corresponden a *líneas de entrada/salida*, y las líneas horizontales de color verde y azul representan los umbrales que se deben superar para que la marca sea reconocida como *discontinua* o de *entrada/salida*, respectivamente.

El algoritmo ha mostrado una buena respuesta en autovías bien conservadas. Las clasificaciones erróneas son ocasionales y se producen fundamentalmente en dos casos. En el primero, la clasificación falla porque la carretera presenta una fuerte curvatura vertical, de modo que al desviarse de la hipótesis de mundo plano las líneas viales dejan de ser paralelas en la imagen de perspectiva inversa y el carril no puede ajustarse correctamente a ambos lados simultáneamente, provocando que el perfil de alguna línea no se extraiga correctamente (figura 5.14). En el segundo caso el fallo se debe a defectos de pintura en el asfalto (figura 5.15).

Pueden verse más ejemplos en la figura 4.33, dentro del apartado 4.5.2. Al estar presentada en blanco y negro, los tipos de línea vial se señalan variando el grosor y espaciado de los puntos.

5.2. Sistema de Alerta por Salida de Carril

Esta es la última etapa del sistema completo. Haciendo uso de todos los algoritmos anteriores, se crea un sistema que generará las alertas necesarias para intentar prevenir una salida no intencionada del carril.

5.2.1. Reconocimiento automático del número de carriles

El sistema de detección y seguimiento de carriles utiliza la clasificación de líneas viales para rastrear, en caso necesario, otras líneas y detectar así de forma automática el número de carriles presentes en un momento determinado, como ya se adelantó en el capítulo 4. Primero se detectan y clasifican los dos bordes del carril por el que se circula. Cuando un borde se reconoce como línea discontinua o de entrada/salida, se asume la existencia de un carril hacia ese mismo lado. En ese caso se extrae el perfil del nuevo borde de

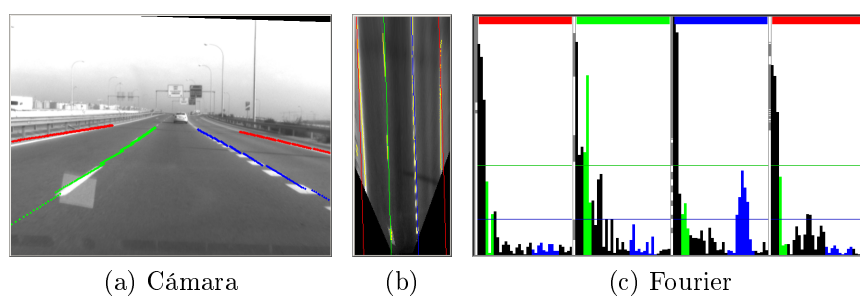


Figura 5.12: Ejemplo de clasificación con tres carriles

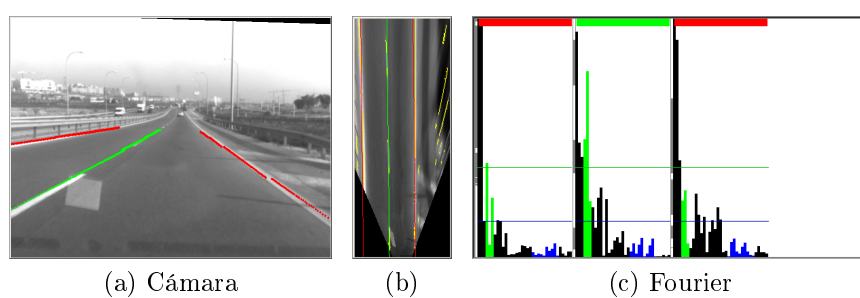


Figura 5.13: Ejemplo de clasificación con dos carriles.

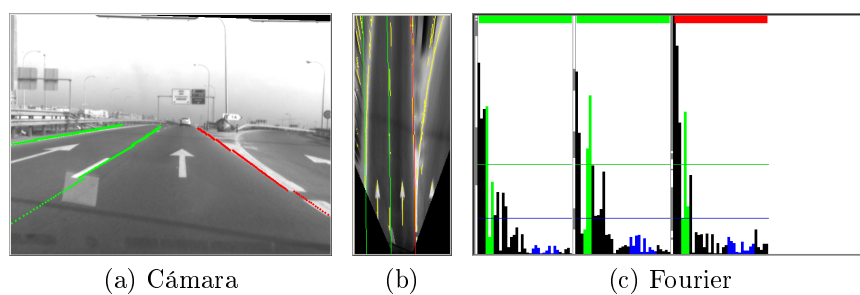


Figura 5.14: Ejemplo de clasificación errónea provocada por la curvatura vertical.

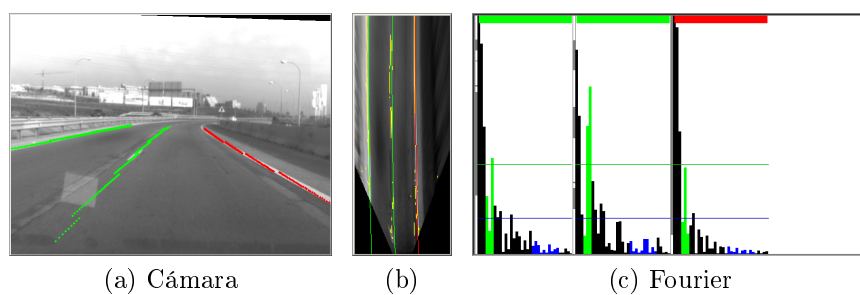


Figura 5.15: Ejemplo de clasificación errónea provocada por defectos de pintura.

carril para su clasificación, asumiendo también que todos los nuevos carriles coinciden en anchura con el carril original. De este modo el sistema es capaz de detectar hasta un total de tres carriles, el propio y los dos adyacentes. La figura 5.16 ilustra algunos ejemplos de los tipos de carretera que pueden ser reconocidos por este sistema.

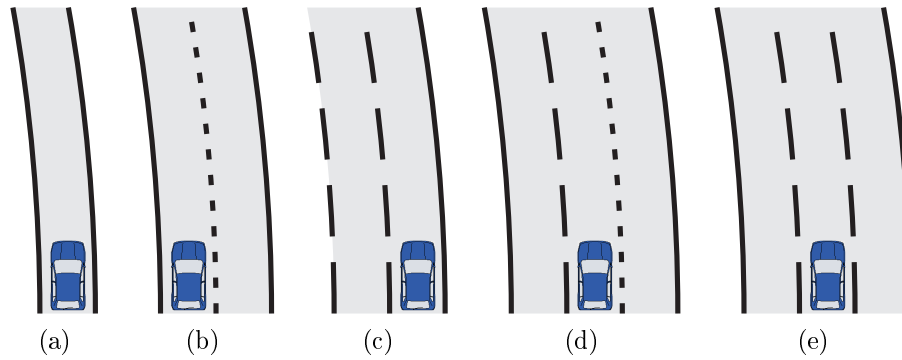


Figura 5.16: Algunos ejemplos de carretera.

5.2.2. Generación de alertas

Hasta ahora el sistema es capaz de reconocer de la carretera hasta tres carriles con su perfil y posición relativa al vehículo, pudiéndose medir en todo momento la distancia a cada línea vial.

Pero un sistema de Alerta por salida de carril debe reconocer además cuándo una maniobra es intencionada o no. Una manera de hacerlo es suponer que cuando el conductor está accionando los intermitentes o el freno significa que es consciente de la maniobra que está realizando. Sin embargo, que el conductor sea consciente no significa necesariamente que la maniobra no entrañe peligro. Por eso, el sistema que se propone en esta tesis utiliza también la información del tipo de línea vial para decidir la conveniencia o no de una alerta, preocupándose de si la maniobra, aunque intencionada, está señalizada correctamente.

Para conocer el estado de los intermitentes, el IVVI se ha equipado con un sistema que permite leer el estado de las luces del vehículo (incluyendo las de freno y cruce). Esto, junto a los algoritmo de detección y seguimiento de carriles y de clasificación de líneas viales, permite al Sistema de Alerta por Salida de Carril calcular la distancia a los bordes del carril, y avisar al conductor, bien cuando está realizando una maniobra no señalizada (no ha activado el intermitente), o bien cuando está cruzando una línea continua (independientemente del estado de los intermitentes).

La predicción de la salida de carril se ha implementado a partir del cálculo del T.L.C. (*“Time to Line Crossing”*, o “tiempo para el cruce de la línea”), tal como se ilustra en la figura 5.17. A partir de la medición de la distancia lateral, y utilizando la información de 8 instantes anteriores, se ajusta un

polinomio de segundo orden para predecir la trayectoria del vehículo respecto del carril. Cuando la trayectoria intersecta con alguno de los límites d_{max} o d_{min} antes de 12 fotogramas (aproximadamente un segundo), el sistema considerará que se está iniciando una maniobra de cambio de carril. Las distancias d_{max} o d_{min} son los valores de la distancia lateral para los cuales las ruedas del vehículo comienzan a pisar una línea vial, y vienen dadas por:

$$\begin{cases} d_{\text{máx}} = \frac{W}{2} - \left(d + \frac{w}{2}\right) \\ d_{\text{mín}} = \left(d + \frac{w}{2}\right) - \frac{W}{2} \end{cases} \quad (5.2)$$

donde W es el ancho del carril, y w es el ancho del vehículo.

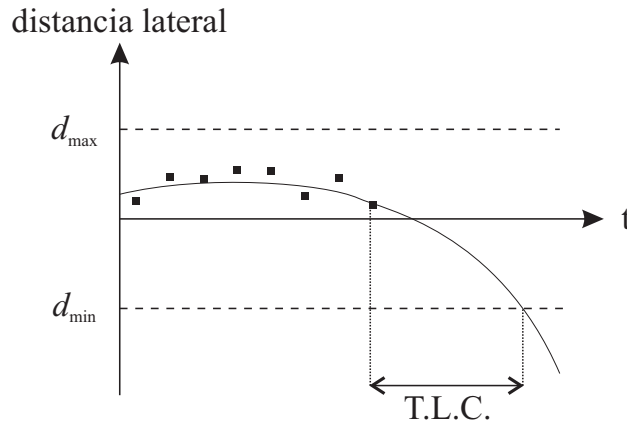


Figura 5.17: Predicción del cruce de la línea.

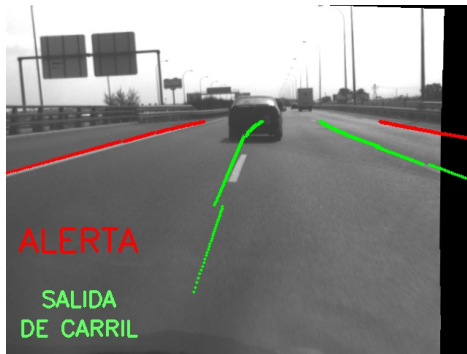
Una vez detectada la maniobra de cambio de carril, el sistema lanza una alerta cuando se produce alguno de los siguientes casos:

- El vehículo va a cruzar una línea discontinua con los intermitentes apagados.
- El vehículo va a cruzar una línea continua, con independencia del estado de los intermitentes.

5.2.3. Resultados

La figura 5.18 muestra varios ejemplos en los que se alerta de salidas de carril no intencionadas. En general el comportamiento del sistema es adecuado, y todos los cambios de carril fueron detectados.

Sin embargo, se producen ocasionalmente falsos positivos, en especial en presencia de lo que algunos autores denominan “singularidades” de la carretera, como por ejemplo las bifurcaciones, incorporaciones y salidas de la vía, y en general, en aquellos casos en los que hay algún carril que va variando su anchura de forma progresiva. En estos casos el algoritmo de seguimiento



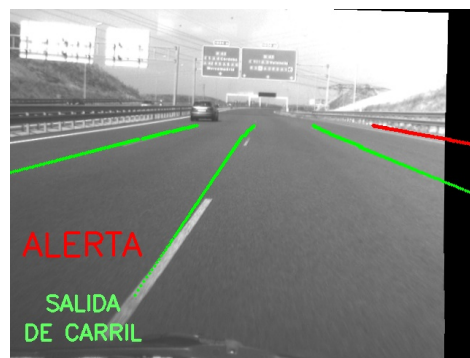
(a) Salida de carril a la izquierda (cruce de línea discontinua)



(b) Salida de carril a la derecha (cruce de línea discontinua)



(c) Salida de carril a la derecha (cruce de línea continua)



(d) Salida de carril a la izquierda (cruce de línea discontinua)



(e) Salida de carril a la izquierda (cruce de línea discontinua)



(f) Salida de carril a la derecha (cruce de línea continua)

Figura 5.18: Ejemplos de alertas por salida de carril

tiende a fijarse en uno de los bordes, y durante unos pocos fotogramas interpreta que el carril se desvía de forma brusca hacia un lateral. La figura 5.19 muestra varios ejemplos de este caso. Otro tipo de fallo lo provocan las falsas clasificaciones, pudiéndose generar alertas por cruce de línea continua cuando ésta es discontinua, aunque este fallo suele ser de corta duración, ya que en condiciones normales las falsas clasificaciones no duran más de dos fotogramas.



(a) Falsa alerta por salida a la izquierda, al pasar por una incorporación.



(b) Falsa alerta por salida a la izquierda, al incorporarse a la autovía.



(c) Falsa alerta por salida a la izquierda, al pasar por una incorporación.



(d) Alerta por cruce de línea continua en lugar de discontinua.

Figura 5.19: Ejemplos de falsas alertas. La mayoría de las falsas alertas se producen en salidas o incorporaciones a la autovía (casos (a), (b) y (c)), porque el carril tiende a seguir la línea continua de la derecha. Las alertas erróneas se producen por una mala clasificación de la línea vial (en el caso (d) la provoca un vehículo que oculta un tramo de la línea).

Capítulo 6

Conclusiones y Trabajos Futuros

Comenzaba el capítulo 1 describiendo cómo las crecientes necesidades de transporte no dejan de hacer crecer el parque automovilístico y el número de desplazamientos. También se explicaba, a través de las estadísticas de accidentes de tráfico, cómo la tecnología del automóvil tiene como contrapartida un considerable riesgo para la salud humana, que es el precio que hoy por hoy hay que pagar para conseguir desplazamientos cada vez más rápidos y a mayor distancia. En este ámbito se enmarca esta tesis, con la que se pretende avanzar en una tecnología que minimice estos riesgos.

A juzgar por las estadísticas, la falta de atención y la conducción temeraria parecen estar detrás de muchos de estos accidentes. Es por ello que se planteó como objetivo de esta tesis el desarrollo de un sistema de ayuda la conducción basado el análisis de imágenes, bajo los siguientes criterios:

- Que sea capaz de detectar la realización de maniobras arriesgadas o incorrectas, en concreto las relacionadas con las salidas o cambios de carril que se realizan de forma no intencionada, o que aún siendo intencionadas incumplen la normativa o implican un riesgo de accidente.
- Que sea de fácil instalación en el vehículo.
- Que suponga un coste asequible dentro del sector del automóvil.

El desarrollo de este sistema ha dado lugar, a su vez, a la implementación de cinco algoritmos:

- Sistema de autocalibración.
- Sistema de detección y seguimiento de carriles rectos con el filtro de Kalman (finalmente descartado).
- Sistema de detección y seguimiento de carriles curvos con el filtro de partículas.
- Sistema de clasificación de líneas viales.
- Sistema de alerta por salida de carril.

6.1. Aportaciones de la tesis

Los algoritmos anteriores son la respuesta que se ha dado a una serie de problemas que ha sido necesario resolver durante el desarrollo de esta tesis. Éstos ya se mencionaban en el primer capítulo, y se relacionan a continuación junto a la solución implementada.

Objetivo	Solución implementada
a) Reconstrucción del perfil de la carretera.	Se ha definido un modelo que aproxima la curvatura horizontal con una parábola.
b) Determinación de la posición del vehículo dentro de la vía.	Se puede medir la distancia a ambos lados del carril.
c) Identificación de los distintos tipos de líneas viales.	El sistema es capaz de distinguir entre tres clases de líneas viales (continuas, discontinuas de separación de carriles normales y discontinuas de separación de carriles de entrada/salida).
d) Identificación del tipo de carretera (autopista, autovía, secundaria, etcétera).	El sistema es capaz de distinguir entre carreteras de 1, 2 y 3 carriles.
e) Estimación de la fiabilidad de las observaciones, para poder detectar de forma automática en qué casos el sistema no es capaz de interpretar las observaciones.	En el apartado 4.5.1 se describe la función que se ha utilizado para este propósito.
f) Alertar al conductor ante maniobras incorrectas o peligrosas.	El sistema es capaz de alertar al conductor cuando éste cruza una línea continua, o una línea discontinua sin señalizar la maniobra.
g) Autocalibración del sistema, sin necesidad de recurrir a patrones artificiales.	El sistema puede autocalibrarse en tramos rectos y con líneas viales visibles.

Salvo las dos primeras, estas capacidades desarrolladas son aportaciones originales en el ámbito de los sistemas de asistencia a la conducción.

6.2. Trabajos futuros

El trabajo realizado en esta tesis se ha concretado en una primera versión de un sistema de alerta por salida de carril, pero al mismo tiempo ha abierto también una serie de campos de mejora. Las propuestas de futuro de mayor envergadura pueden agruparse en las tres siguientes líneas de investigación:

6.2.1. Mejoras en el Modelo de Carretera

Mejoras en el Modelo Espacial

El modelo espacial utilizado es válido en la mayoría de las situaciones que se encuentran en la carretera, pero puede ser mejorado de forma que pueda contemplar “singularidades”, como la aparición y desaparición de carriles laterales, y bifurcaciones. Se propone investigar la manera de:

- Incluir en el modelo la curvatura vertical de la carretera. Esto evitaría el problema de las vibraciones del sistema de visión (cf. apartado 4.5.2), mejoraría la clasificación de líneas viales (cf. apartado 5.1.4), y permitiría analizar una mayor distancia de carretera.
- Contemplar la aparición o desaparición progresiva de carriles, así como las bifurcaciones de carriles en dos.
- Ampliar el rango de tipos de marcas viales detectadas. En especial, la posibilidad de reconocer líneas dobles y distinguir entre más tipos de líneas discontinuas facilitaría la diferenciación entre autovías y carreteras convencionales.

Mejoras en el Modelo Temporal

El modelo temporal es el que ofrece la capacidad para predecir la trayectoria del vehículo, además de facilitar el seguimiento del carril en la secuencia de vídeo. El modelo implementado utiliza exclusivamente la información procedente de las cámaras, pero recientemente se ha instalado en IVVI un sistema de navegación que proporciona información inercial del vehículo. Por lo tanto, se propone:

- Incluir información inercial del vehículo en el modelo temporal. Es de suponer que esto mejorará el rendimiento del seguimiento de carriles y permitirá predecir la trayectoria de tal forma que pueda anticiparse la salida del carril con más precisión.

6.2.2. Mejoras en el Sistema de Adquisición de Imágenes

El principal problema que presenta el actual sistema de adquisición de imágenes es la transmisión de vibraciones a las cámaras debido al mecanismo de anclaje al parabrisas (figura 4.31). Por lo tanto, la propuesta en este apartado es:

- Limitar las vibraciones transmitidas al sistema de visión, ya sea de forma mecánica (con un sistema de anclaje alternativo), o mediante un software de estabilización de imágenes.

6.2.3. Integración de la Detección de Carreteras con el Resto de Módulos de Detección del IVVI

En el IVVI se están desarrollando otros módulos de detección: detección de vehículos, detección de señales de tráfico, detección de peatones y detección de la cara del conductor. De la interacción entre todas las capacidades perceptivas del IVVI se pueden obtener, a su vez nuevas capacidades y mejorar las ya existentes. En este sentido, se propone dar al módulo de detección de carreteras la posibilidad de:

- Comunicar al resto de módulos la información obtenida por la detección de carreteras, permitiéndoles así delimitar la posición y el tamaño en la imagen de los vehículos, las señales de tráfico, y los peatones. En concreto, los vehículos están situados sobre la calzada, y las señales de tráfico en los bordes de la misma. Tanto vehículos, como señales y peatones deben tener un tamaño acorde a su distancia, que se puede deducir de su posición en la imagen si se conoce el perfil de la carretera.
- Recibir la información obtenida por otros módulos de detección. Las señales de tráfico suelen avisar con antelación de la presencia de bifurcaciones, incorporaciones, salidas, y la aparición y desaparición de carriles. Por otro lado, la detección de vehículos se puede utilizar para excluir del procesamiento aquellas regiones de la imagen ocupadas por ellos.
- Definir un índice de probabilidad de tipo de carretera (autovía o carretera convencional). Conociendo el número y tipo de carriles, la información dada por las señales de tráfico, y la posición y velocidad relativa de otros vehículos, puede definirse un índice de probabilidad que permita reconocer con una alta fiabilidad por qué tipo de carretera se está circulando.

Apéndice A

Plataforma de Experimentación

A.1. El Vehículo Ivvi

Ivvi (figura A.1) es una plataforma de investigación para la implementación de sistemas basados en visión por computador y que sirvan de ayuda a la conducción. En la actualidad se compone de tres subsistemas:

1. **Sistema de percepción.** Está constituido por:

- *Sensores de percepción del entorno:*
 - Una cámara color (Hitachi KP-D20A) para la detección y análisis de la señalización vertical de la carretera, situada en la parte frontal del vehículo (figura A.1d).
 - Un sistema estéreo formado por dos cámaras CCD blanco y negro de escaneo progresivo (Hitachi KP-F3) para la detección de obstáculos, de peatones (diurna), y de los carriles de la carretera, situado también en la parte frontal (figura A.1c).
 - Una cámara infrarroja (Indigo Omega), para la detección nocturna de peatones (figura A.1e).
- *Sensores de propiocepción* (referidos al sistema conductor-vehículo):
 - Una cámara color para monitorizar la cara del conductor (figura A.1b).
 - Una antena GPS para la medición de la velocidad y posición del vehículo (figura A.1f).
 - Un sistema de monitorización de la conducción, implementado en la universidad, para la lectura del estado de los faros del vehículo (faros delanteros, freno e intermitentes).

2. **Sistema de procesamiento.** Está formado por (figura A.1f):

- Dos PCs con procesador Intel Pentium D 2Ghz y 2GB de memoria DRAM.

- Una tarjeta de adquisición de imágenes Matrox Morphis, a la que se conectan las cámaras infrarroja, a color y de monitorización del conductor.
 - Una tarjeta de adquisición de imágenes Matrox Meteor II Multi-Channel, a la que se conecta el sistema estéreo
 - Una PDA, que hace de interfaz entre el GPS y los PCs.
 - Un router WiFi.
3. **Interfaz con el conductor.** Interacciona con el conductor avisándole de forma acústica sobre la información recogida sobre la carretera y la conveniencia o no de las maniobras que realiza. Actualmente utiliza los altavoces del sistema de sonido del coche (figura A.2b).

En el vehículo IVVI se están desarrollando las siguientes capacidades sensoriales:

- **Detección de la señalización vertical.** Detección e interpretación de las señales de tráfico y los paneles informativos.
- **Detección de vehículos y detección diurna y nocturna de peatones.** Detección y seguimiento de los diversos objetos que rodean al sistema estimando su velocidad y trayectoria.
- **Detección de la carretera.** Se detectarán los diversos carriles que tiene la carretera.
- **Detección de la cara del conductor.** Monitoriza el estado de los ojos y la boca del conductor.

La combinación de estas habilidades dará lugar a un análisis más complejo del entorno, a través de los siguientes cuatro módulos:

- **Módulo anti-colisión.** Tendrá en cuenta la posición del vehículo respecto a las líneas laterales de la carretera y los vehículos que le rodean.
- **Sistema de supervisión de la velocidad.** Indicará al conductor la velocidad correcta en función de la propia velocidad del vehículo, la de los que lo rodean y de las señales viarias.
- **Sistema de alerta por salida de carril.** Alertará al conductor en el caso de que la trayectoria del vehículo le lleve a salirse de la calzada, o hacia un carril que no debe invadirse.
- **Sistema de supervisión del conductor.** Estudia la cara del conductor y le alerta en el caso de que entre en estado de somnolencia.



(a) Vista exterior



(b) Cámara orientada al conductor



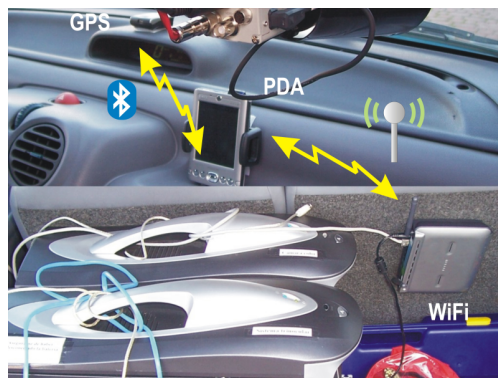
(c) Sistema estéreo



(d) Cámara a color



(e) Cámara infrarroja



(f) Sistema de procesamiento

Figura A.1: Ivvi



(a) Terminal del sistema de procesamiento

(b) Sistema de sonido

Figura A.2: Interfaces del IVVI.

A.2. Capacidades Sensoriales

A.2.1. Detección de Peatones

Según los estudios de accidentes de tráfico en Europa se produce un gran número de accidentes entre peatones y vehículos. Cada año en la Unión Europea más de 200.000 son atropellados, de los cuales más de 6.000 son víctimas mortales. Los atropellos constituyen el 12 % del total de accidentes de carretera, y el 15 % de las muertes en carretera. En concreto, en España los atropellos a peatones constituyen, en entornos urbanos, el segundo tipo de accidente más frecuente (18 %), y el que más víctimas mortales se cobra (37 %) (ver figura 1.2 en la página 3).

En el IVVI se han desarrollado dos sistemas de detección de peatones, uno para detección diurna (figura A.3), y otro para detección nocturna (figura A.4). Una descripción de estos sistemas puede encontrarse en (Bertozzi y otros, 2007; Hilario y otros, 2005b,e,f).

A.2.2. Detección de Señales de Tráfico

La detección automática de señales de tráfico ha recibido un interés creciente por parte de los laboratorios de investigación debido a las aplicaciones que se podrían desarrollar como:

- **Mantenimiento de autopistas.** Actualmente es un operador el que tiene que observar una cinta de vídeo para determinar si la señal de tráfico está en buen estado y goza de buena visibilidad.
- **Inventario de señales en ciudades.** En este entorno las señales no

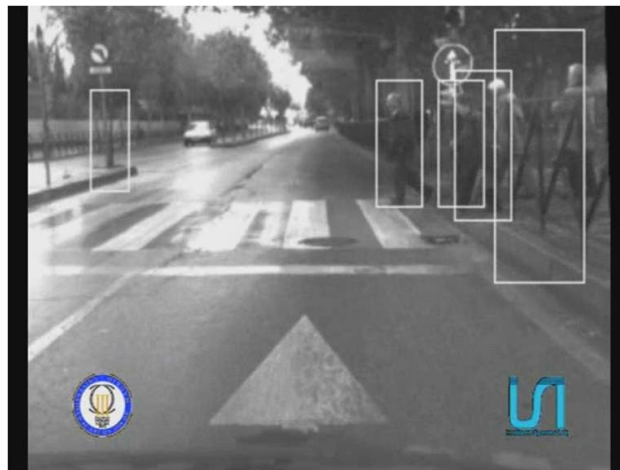


Figura A.3: Detección diurna de peatones.

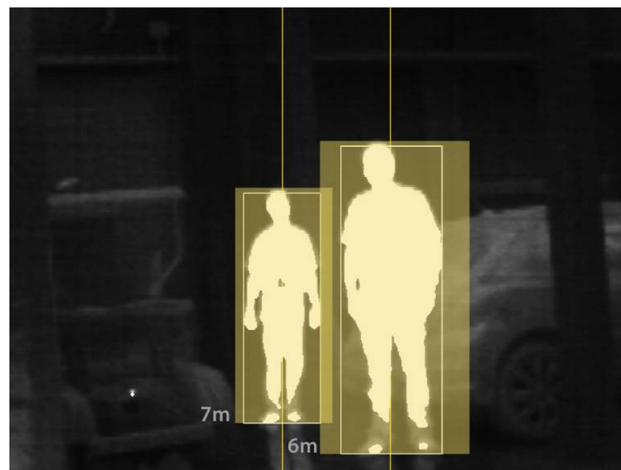


Figura A.4: Detección nocturna de peatones.

están siempre perpendiculares al movimiento del vehículo, hay objetos con el mismo color y las oclusiones son más frecuentes.

- **Sistemas de ayuda a la conducción.** Su interpretación facilita la supervisión de la velocidad, y de la trayectoria del vehículo.

Las principales dificultades a las que se enfrenta la detección de señales son las condiciones de iluminación cambiantes y no controlables, la presencia de otros objetos que dan lugar a oclusiones y sombras, y por último el rango de posibles variaciones de la apariencia del objeto en la imagen es muy grande.

El Sistema de Reconocimiento de Señales de Tráfico del IVVI mide la velocidad del vehículo mediante una antena GPS, y reconoce mediante visión por computador la presencia de señales de peligro, prohibición, ceda el paso, y fin de prohibición (figura A.5). Este sistema emite alertas a través de los altavoces del vehículo en los casos en que la velocidad del vehículo:

- supera el límite establecido por una señal.
- es inferior a la mínima permitida.
- supera un umbral, definido por el usuario, en presencia de señales de peligro o ceda el paso.

Una descripción de los algoritmos de detección de señales de tráfico desarrollados por el Grupo de Sistemas Inteligentes puede encontrarse en (Carrasco y otros, 2008, 2007, 2006; De la Escalera y otros, 2004, 2003; de la Escalera y otros, 2001; De la Escalera y otros, 1997, 1994a,b)



Figura A.5: Sistema de Reconocimiento de Señales de Tráfico

A.2.3. Detección de Vehículos

El módulo de detección de vehículos es fundamental para las siguientes tareas:

- **Seguimiento en pelotón.** Los vehículos circulan en grupo a altas velocidades y con distancias de separación pequeñas.
- **Parar y marchar (“Stop&go”).** Básicamente es lo mismo que el problema anterior pero para el caso de conducción dentro de una ciudad, donde las paradas son más frecuentes.
- **Ángulo muerto.** El sistema tiene que detectar que hay otro vehículo aproximándose y alertar al conductor para evitar que éste inicie un adelantamiento.
- **Supervisor de maniobras** propias y de los demás vehículos.

El sistema de Detección de Vehículos del IVVI , reconoce la presencia y posición de otros vehículos mediante visión por computador, combinando el análisis de movimientos con algoritmos genéticos (Collado y otros, 2007; Hilario y otros, 2006, 2005a,c,d; Collado y otros, 2004; Hilario y otros, 2004). La figura A.6 presenta varios ejemplos de detección de vehículos.



Figura A.6: Detección de vehículos

A.2.4. Reconocimiento de la cara del conductor

Una de las principales causas de accidentes en carretera, aunque difícil de medir, es probablemente la falta de atención provocada por un estado de somnolencia.

El Sistema de Reconocimiento de la Cara del Conductor utiliza un enfoque basado en varios algoritmos de aprendizaje para localizar la posición y estado de los ojos del conductor (figura A.7). Partiendo de las mediciones de la frecuencia y latencia del parpadeo del conductor, se define un indicador del estado de somnolencia. El sistema emite una alerta al conductor alcanza un cierto estado de somnolencia (Flores y otros, 2007).

A.2.5. Detección de los Carriles de la Carretera

El desarrollo de esta habilidad constituye el objeto de esta tesis, y se describe en detalle en los próximos capítulos, y en Collado y otros (2008, 2006, 2005b,a).



Figura A.7: Reconocimiento del estado de somnolencia

Apéndice B

Rectificación de imágenes estéreo

B.1. Introducción

Se tiene el sistema estéreo mostrado en la figura B.1, donde se han colocado: Dos cámaras exactamente iguales: idéntico elemento sensor, ganancia. Las ópticas tienen distancias focales idénticas. Están situadas de forma que sus ejes ópticos sean colineales. Están separadas una distancia D (denominada en inglés *baseline*).

El punto del mundo 3D $\mathbf{M} = (x, y, z)^T$ se proyecta en las dos imágenes en los píxeles $\mathbf{m}_i = (u_i, v_i)^T$ y $\mathbf{m}_d = (u_d, v_d)^T$. En el caso general las coordenadas de estos dos puntos no serán las mismas, denominándose *disparidad* a la distancia o diferencia entre ambas.

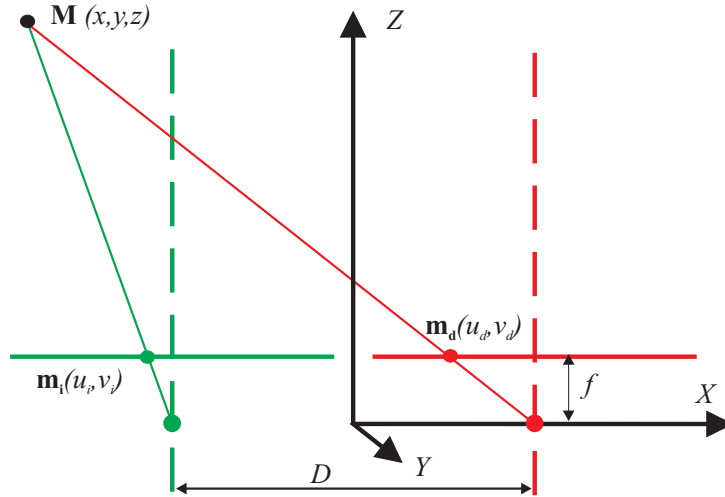


Figura B.1: Disposición “ideal” de un sistema estéreo.

Según el modelo de óptica *pin-hole* y tomando como origen de referencia

el punto medio entre las cámaras, las ecuaciones para cada una de ellas son:

$$\begin{aligned}\frac{u_i}{f} &= \frac{x + D/2}{z} \\ \frac{u_d}{f} &= \frac{x - D/2}{z}\end{aligned}\tag{B.1}$$

donde x y z son las incógnitas y el resto datos. Si se igualan las expresiones para la coordenada horizontal x :

$$x = z \frac{u_i}{f} - \frac{D}{2} = z \frac{u_d}{f} + \frac{D}{2}\tag{B.2}$$

$$z = f \frac{D}{u_i - u_d} = f \frac{D}{d}\tag{B.3}$$

Siendo d la disparidad, que en el caso particular de que ser las cámaras paralelas será siempre positiva.

La coordenada y será:

$$y = \frac{v_i}{f} z = \frac{v_i}{f} \frac{Df}{d} = \frac{Dv_i}{d}\tag{B.4}$$

Si se compara con el resultado obtenido con la otra imagen:

$$\frac{v_d}{f} = \frac{y}{z} = \frac{Dv_i}{d} \frac{d}{Df} = \frac{v_i}{f} \Rightarrow v_i \equiv v_d\tag{B.5}$$

Como son iguales, si se toma un punto de la cámara izquierda, su correspondiente en la imagen derecha estará a la misma altura. En el caso general en el que las cámaras no sean paralelas se sigue cumpliendo que dado un punto en una imagen, los posibles píxeles candidatos en la otra imagen se encuentran en una línea recta, denominada *línea epipolar*, y en el caso visto hasta ahora son siempre líneas paralelas horizontales. Esta simplicidad en la formulación matemática, que además consigue que la búsqueda de puntos semejantes sea un problema 1D en lugar de 2D, se ha basado en una serie de premisas nada fáciles de cumplir en la práctica: que ambas cámaras sean exactamente iguales y que los ejes ópticos sean paralelos. Esto puede verse en la Fig. 2. Para construir el sistema estéreo se han colocado dos cámaras con ópticas del mismo modelo, y se ha procurado que la estructura mecánica asegure que estén paralelas. Puede comprobarse, siguiendo las líneas blancas horizontales, que no lo son totalmente. Aunque pueda pensarse que las diferencias son muy pequeñas, en la Fig. 2-c se tiene el resultado de obtener el mapa de disparidades en el que se observa que los resultados erróneos. Hay que hacer notar que estos errores han sido debidos a que hemos buscado los puntos correspondientes entre las dos imágenes en las líneas epipolares que se suponían serían las líneas horizontales. En el caso general hubiera pasado lo mismo porque el error de fondo está en las tolerancias en la fabricación de las cámaras, las ópticas y la base mecánica. Por tanto el problema de la

búsqueda de puntos correspondientes sigue siendo 2D a menos que se realice un preprocesamiento de las imágenes estéreo.

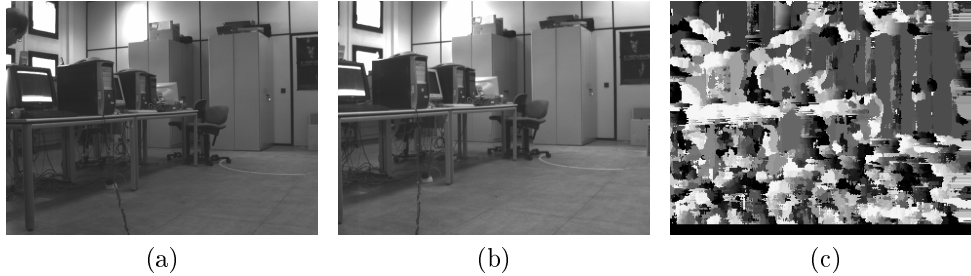


Figura B.2: Necesidad de procesar las imágenes de un sistema estéreo; (a) y (b) Imágenes izquierda y derecha originales; (c) Mapa de disparidades entre (a) y (b).

La rectificación (figura B.3) es la transformación de una pareja de sistemas ópticos en otros dos que sean colineales y donde las líneas epipolares sean paralelas a unos de los ejes de las imágenes. Se puede pensar en este proceso como si se hubiera rotado a las cámaras originales alrededor de sus centros ópticos. En los siguientes apartados se va a ver cómo puede realizarse esto en la práctica y qué pasos hay que dar.

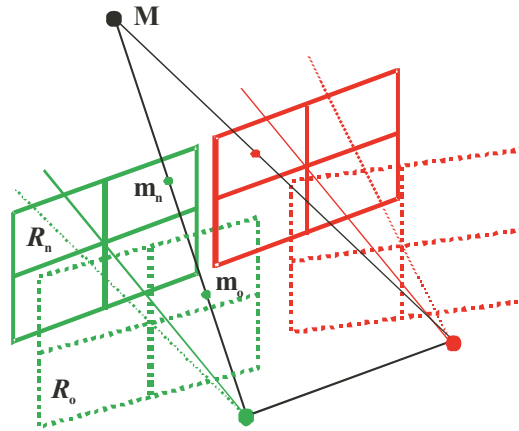


Figura B.3: Rectificación

B.2. Modelo de cámara en coordenadas homogéneas

Siguiendo la notación propuesta por Ayache en (Ayache, 1991), el modelo de la cámara (figura B.4) consiste en el centro óptico, \mathbf{c} , el plano de la imagen (también denominado plano de la retina), R . Un punto del mundo 3D, $\mathbf{M} = (\mathbf{xyz})^T$ en coordenadas del mundo, se proyecta en la imagen en un punto

2D con coordenadas $\mathbf{m} = (\mathbf{u}, \mathbf{v})^T$, donde \mathbf{m} es la intersección del plano R con la línea de contiene a \mathbf{M} y a \mathbf{c} . Esto mismo, expresado en coordenadas homogéneas es:

$$\tilde{\mathbf{m}} = (U, V, S)^T \quad (\text{B.6})$$

con:

$$u = \frac{U}{S} \quad (\text{B.7})$$

donde los puntos para los que S vale cero pertenecen al plano focal F . De esta forma, utilizando este sistema de coordenadas, estos puntos ya no son una excepción al valer infinito, sino que corresponden a aquellos puntos en los que S vale cero.

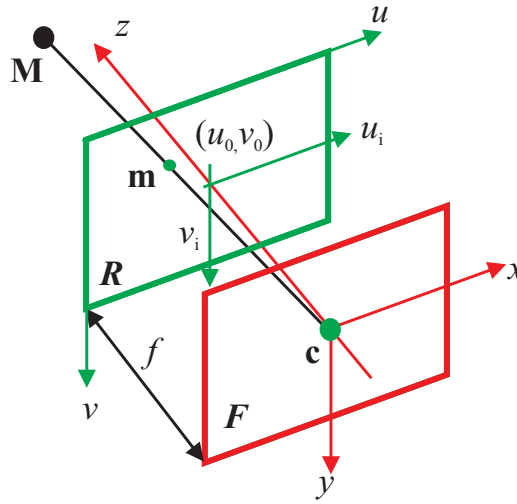


Figura B.4: Modelo de cámara.

De forma análoga:

$$\tilde{\mathbf{M}} = (x, y, z, 1)^T \quad (\text{B.8})$$

y la matriz de proyección de perspectiva que relaciona ambos puntos es:

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{M}} \quad (\text{B.9})$$

En el siguiente apartado se analizará la matriz de proyección con más detalle.

B.3. Matriz de proyección de perspectiva y geometría epipolar

La matriz de proyección de perspectiva vista antes puede descomponerse en el producto de otras dos matrices:

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{M}} = \mathbf{A}\mathbf{G}\tilde{\mathbf{M}} \quad (\text{B.10})$$

donde \mathbf{A} representa los parámetros intrínsecos de la cámara y \mathbf{G} los extrínsecos.

Para los parámetros intrínsecos se tiene que:

$$\frac{u_i}{f} = \frac{x}{z} \frac{v_i}{f} = \frac{y}{z}$$

$$\begin{pmatrix} U \\ V \\ S \end{pmatrix} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (\text{B.11})$$

Hay un nuevo cambio de sistema de coordenadas ya que un punto en la imagen está medido en píxeles, no en milímetros, y su origen es la esquina superior izquierda, no el punto de corte del eje óptico (figura B.4). Por ello se debe de conocer el paso de milímetros a píxeles para cada eje, (k_u, k_v) , y el las coordenadas del corte del eje óptico con el plano de la imagen, el punto principal, respecto a la esquina superior izquierda (u_o, v_o) , en píxeles.

$$\begin{aligned} u &= k_u u_i + u_0 \\ v &= k_v v_i + v_0 \end{aligned} \quad (\text{B.12})$$

Despejando:

$$\begin{aligned} u &= k_u f \frac{x}{z} + u_0 = \frac{f_u x + u_0 z}{z} \\ v &= k_v f \frac{y}{z} + v_0 = \frac{f_v y + v_0 z}{z} \end{aligned} \quad (\text{B.13})$$

Por lo que se obtiene:

$$\mathbf{A} = \begin{pmatrix} f_u & 0 & u_o \\ 0 & f_v & v_o \\ 0 & 0 & 1 \end{pmatrix} \quad (\text{B.14})$$

La matriz \mathbf{G} está compuesta por una matriz de rotación, \mathbf{R} , de dimensiones 3x3 y un vector de traslación, \mathbf{t} , que representan la posición y orientación de la cámara respecto al sistema de coordenadas del mundo.

$$\mathbf{G} = (\mathbf{R}|\mathbf{t}) \quad (\text{B.15})$$

Por lo que:

$$\tilde{\mathbf{m}} = \begin{pmatrix} U \\ V \\ S \end{pmatrix} = \begin{pmatrix} f_u & 0 & u_o \\ 0 & f_v & v_o \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{00} & r_{01} & r_{02} & t_x \\ r_{10} & r_{11} & r_{12} & t_y \\ r_{20} & r_{21} & r_{22} & t_z \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \tilde{\mathbf{P}}\tilde{\mathbf{M}} \quad (\text{B.16})$$

De esta matriz se puede obtener la posición del centro óptico, que luego servirá para la rectificación. Para ello se ve que puede reescribirse como:

$$\tilde{\mathbf{m}} = \begin{pmatrix} U \\ V \\ S \end{pmatrix} = \tilde{\mathbf{P}}\tilde{\mathbf{M}} = \begin{pmatrix} \mathbf{q}_1^T | q_{14} \\ \mathbf{q}_2^T | q_{24} \\ \mathbf{q}_3^T | q_{34} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = (\mathbf{P} | \tilde{\mathbf{p}}) \begin{pmatrix} \mathbf{M} \\ 1 \end{pmatrix} \quad (\text{B.17})$$

¿Qué significan cada una de las filas de la matriz? Se comentó antes que el plano para el que S es cero, $\mathbf{q}_3^T \mathbf{M} + q_{34} = 0$, es el plano focal, y los planos $\mathbf{q}_1^T \mathbf{M} + q_{14} = 0$ y $\mathbf{q}_2^T \mathbf{M} + q_{24} = 0$ intersectan al plano de la imagen formando el eje vertical (u nulo) y horizontal (v nulo) respectivamente. El centro óptico, \mathbf{c} , será el punto en el que intersecten los tres planos, luego debe cumplir:

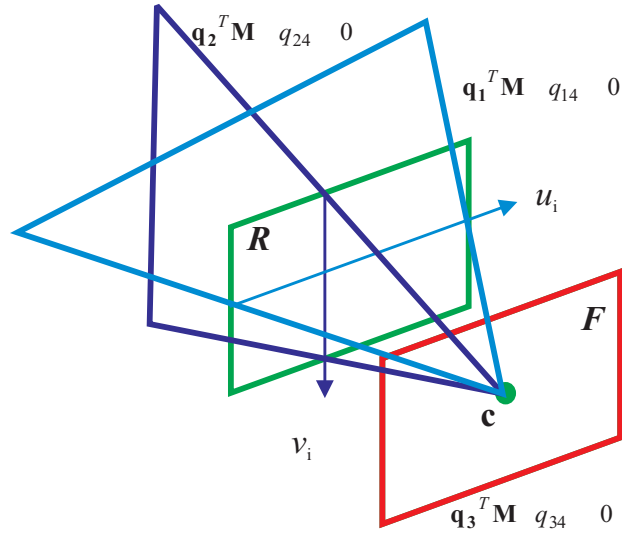


Figura B.5: Interpretación de las filas de la matriz de proyección de perspectiva.

$$\tilde{\mathbf{P}} \begin{pmatrix} \mathbf{c} \\ 1 \end{pmatrix} = \mathbf{0} \Rightarrow (\mathbf{P} | \tilde{\mathbf{p}}) \begin{pmatrix} \mathbf{c} \\ 1 \end{pmatrix} = \mathbf{0} \Rightarrow \mathbf{P}\mathbf{c} + \tilde{\mathbf{p}} = \mathbf{0} \Rightarrow \begin{cases} \mathbf{c} = -\mathbf{P}^{-1}\tilde{\mathbf{p}} \\ \tilde{\mathbf{p}} = -\mathbf{P}\mathbf{c} \end{cases} \quad (\text{B.18})$$

Y entonces:

$$\tilde{\mathbf{P}} = (\mathbf{P} | -\mathbf{P}\mathbf{c}) \quad (\text{B.19})$$

Debido a que un rayo óptico asociado con un punto \mathbf{m} en la imagen es la línea que une ese punto con el centro óptico, se deriva que:

$$\tilde{\mathbf{m}} = \tilde{\mathbf{P}}\tilde{\mathbf{M}} = (\mathbf{P}|\tilde{\mathbf{p}}) \begin{pmatrix} \mathbf{M} \\ 1 \end{pmatrix} = (\mathbf{P}|- \mathbf{P}\mathbf{c}) \begin{pmatrix} \mathbf{M} \\ 1 \end{pmatrix} = \mathbf{P}(\mathbf{M} - \mathbf{c}) \quad (\text{B.20})$$

$$\mathbf{M} = \mathbf{c} + \lambda \mathbf{P}^{-1} \tilde{\mathbf{m}}$$

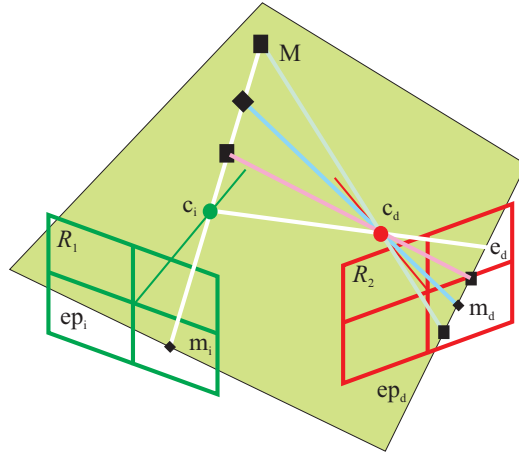


Figura B.6: La línea epipolar.

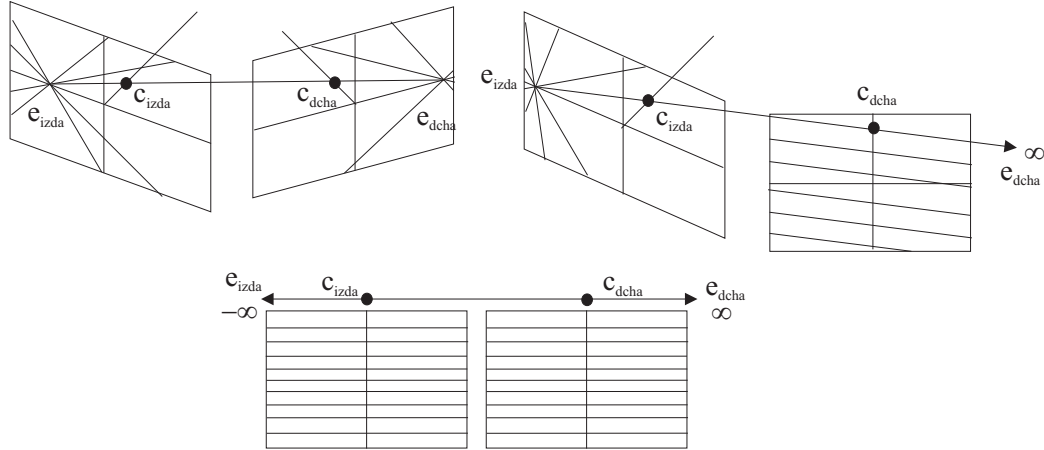


Figura B.7: Posición de las líneas epipolares.

Se comentó al principio del documento que si se tienen las coordenadas de las proyecciones de un punto del mundo sobre dos imágenes, es fácil obtener las coordenadas tridimensionales. Pero, conociendo la proyección en una imagen, ¿Cómo encontrar el píxel de la segunda? La geometría epipolar antes enunciada es una primera ayuda al no resultar necesario buscar la segunda proyección en toda la imagen ya que se encontrará en la *línea epipolar*. Así,

para generar la proyección del punto del mundo \mathbf{M} se traza una línea que lo una al foco de la imagen izquierda. Si sólo se dispusiese de esa imagen no se podría calcular la distancia a la que se encuentra el objeto, pero sí su dirección. Si se dibujasen todas las proyecciones de los posibles puntos que se encuentran en esa recta se obtendría la línea epipolar. Ésta se define por tanto, como el corte en el plano de la imagen de una de las cámaras del plano que contiene a los dos puntos focales y el punto en el espacio. Todas las líneas epipolares coinciden en el mismo punto (epipolo). El epipolo de una cámara es la proyección del centro óptico de la otra, ya que es el único punto común a todas las rectas de proyección. Si el valor de su coordenada homogénea S es cero, el epipolo está en el infinito y esto quiere decir que la proyección del centro óptico está en el plano focal de la otra cámara (figura B.8) y las líneas epipolares serán paralelas. Si además la línea que une ambos centros ópticos es paralela a uno de los ejes de la imagen serán horizontales o verticales.

B.4. Calibración del sistema estéreo

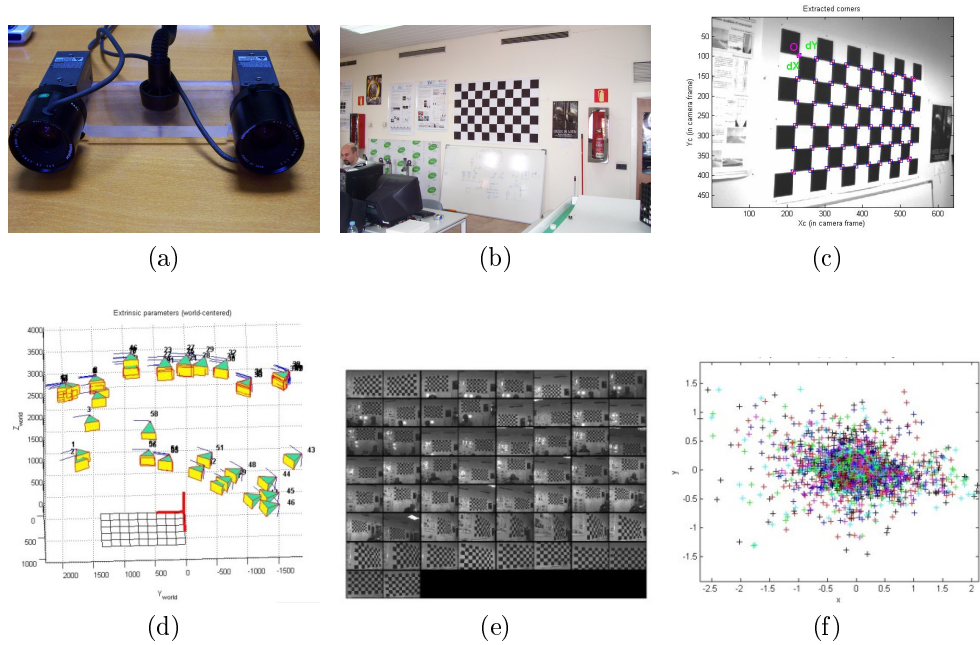


Figura B.8: Calibración del sistema estéreo; (a) sistema; (b) y (c) patrón de calibración; (d) posiciones de la cámara respecto al patrón; (e) imágenes tomadas con la cámara izquierda; (f) errores.

El método de rectificación implementado sigue el desarrollado por Fusiello (Fusiello y otros, 2000), que supone que las matrices de proyección son conocidas. Se pueden consultar otras referencias (Hartley y Gupta, 1993; Robert y otros, 1997) para el caso de que no lo sean. Por ello debe de calibrarse el sistema estéreo (figura 7-a) y obtener los parámetros intrínsecos

(distancia focal, centro de la imagen, distorsiones) y extrínsecos (matriz de rotación y vector de traslación). Para ello se ha utilizado el Toolbox para Matlab desarrollado en el California Institute of Technology (Bouguet, 2006) que se está convirtiendo en el software estándar de calibración, aunque pueden encontrarse métodos alternativos (Papadimitriou y Dennis, 1996). Éste Toolbox implementa (entre otros) el método desarrollado por Zhang (Zhang, 1999, 2000). Requiere un patrón de calibración plano (figura B.8b), y capturar al menos dos imágenes en posiciones distintas que no tiene por qué ser conocidas (figura B.8(d)(e)). Se calibran los parámetros intrínsecos antes mencionados gracias a la homografía (coincidencia gráfica) entre el modelo plano y su imagen. Primero se deja de lado la distorsión mientras se calcula mediante una solución analítica el resto de parámetros, y luego se utiliza un método de optimización no lineal para obtenerla. Se necesita un patrón similar a un tablero de ajedrez donde el usuario señala las cuatro esquinas y el sistema determina el resto de ellas de forma automática. Respecto a la distorsión (figura B.9), tiene en cuenta la radial y también la tangencial (Brown, 1971). Así:

$$\begin{aligned} x_r &= x_i + x_i(k_1 r^2 + k_2 r^4) + 2p_1 x_i y_i + p_2(r^2 + 2x_i^2) \\ y_r &= y_i + y_i(k_1 r^2 + k_2 r^4) + p_1(r^2 + 2y_i^2) + 2p_2 x_i y_i \end{aligned} \quad (\text{B.21})$$

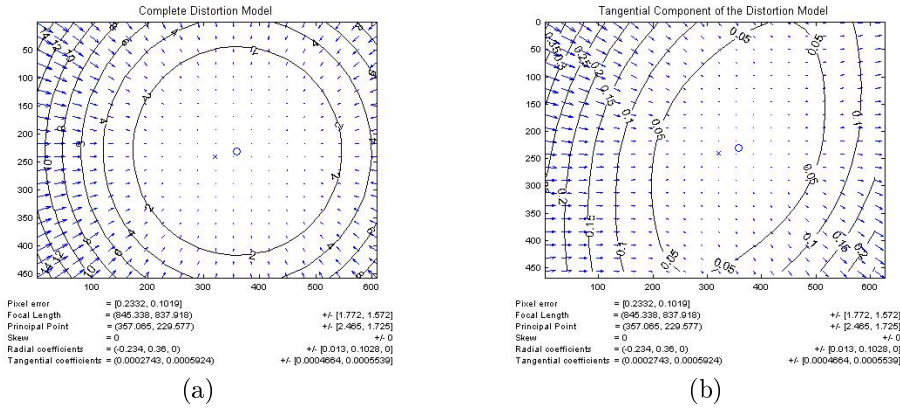


Figura B.9: Distorsión; (a) radial; (b) tangencial

Siendo (x_r, y_r) las coordenadas reales de los píxeles, (x_i, y_i) las coordenadas ideales, k_1 y k_2 los coeficientes de la distorsión radial, p_1 y p_2 los coeficientes de la distorsión tangencial y r la distancia al centro de la imagen.

Si las imágenes han sido tomadas por un sistema estéreo se obtiene también la relación entre las dos cámaras según:

$$\mathbf{M_D} = \mathbf{R}\mathbf{M_I} + \mathbf{t} \quad (\text{B.22})$$

Siendo \mathbf{R} una matriz de rotación y \mathbf{t} el vector de traslación, según los sistema de referencia de la figura B.10.

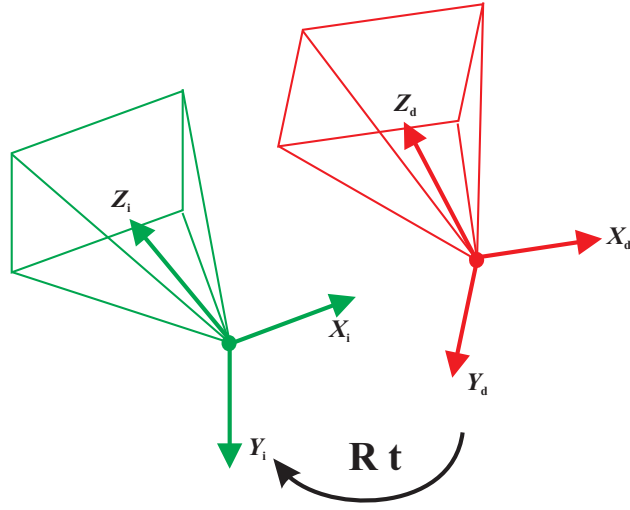


Figura B.10: Patrón de calibración y sistemas de coordenadas

El sistema estéreo del vehículo IVVI se colocó en 58 posiciones distintas (figura B.8(d)(e)) y los resultados obtenidos están reflejados en la tabla B.1.

Tabla B.1: Valores de la calibración

Cámara	k_1	K_2	p_1	p_2	u_o	v_o	f_x	f_y
Izquierda	-0.24386	0.19161	0.00254	-0.00761	308.49	237.40	815.03	806.27
Derecha	-0.25275	0.32041	0.00470	-0.00905	301.74	226.90	815.40	807.86
$\mathbf{R} = \begin{pmatrix} 1,0000 & -0,0047 & 0,0055 \\ 0,0047 & 1,0000 & -0,0025 \\ -0,0055 & 0,0025 & 1,0000 \end{pmatrix}; \quad \mathbf{t} = \begin{pmatrix} -147,86 & 0,96 & 11,3 \end{pmatrix}^T$								

B.5. Rectificación

Como se ha comentado antes, el método de rectificación exige conocer las matrices de rotación y la traslación entre el sistema de coordenadas del mundo y el de cada cámara:

$$\begin{aligned} \mathbf{M}_I &= \mathbf{R}_I \mathbf{M}_w + \mathbf{t}_I \\ \mathbf{M}_D &= \mathbf{R}_D \mathbf{M}_w + \mathbf{t}_D \end{aligned} \quad (\text{B.23})$$

Mediante la ecuación (B.22) se comprueba que:

$$\begin{aligned} \mathbf{R}_I &= \mathbf{I} & \mathbf{t}_I &= \mathbf{0} \\ \mathbf{R}_D &= \mathbf{R} & \mathbf{t}_D &= \mathbf{t} \end{aligned} \quad (\text{B.24})$$

Por lo que las ecuaciones (B.16) y (B.17) se tiene que:

$$\begin{aligned} \tilde{\mathbf{P}}_{oi} &= (\mathbf{P}_{oi} | \tilde{\mathbf{p}}_{oi}) = \mathbf{A}_{oi} (\mathbf{R}_{oi} | \mathbf{t}_{oi}) = (\mathbf{A}_{oi} | \mathbf{0}) \\ \tilde{\mathbf{P}}_{od} &= (\mathbf{P}_{od} | \tilde{\mathbf{p}}_{od}) = \mathbf{A}_{od} (\mathbf{R} | \mathbf{t}) = (\mathbf{A}_{od} \mathbf{R} | \mathbf{A}_{od} \mathbf{t}) \end{aligned} \quad (\text{B.25})$$

El método de rectificación implementado va a definir dos nuevas matrices de perspectiva que cumplan las siguientes condiciones:

- Debe preservar los centros ópticos
- Los planos focales deben de ser coplanares por lo que contienen la recta que une los centros ópticos. Con esta condición los epipolos están en el infinito y las líneas epipolares son paralelas.
- Además de paralelas, se quiere que las líneas epipolares sean horizontales, para ello la recta anterior debe de ser paralela al eje horizontal de ambas cámaras.
- Por último, los puntos conjugados, los de ambas imágenes que corresponde al mismo punto del mundo, deben de tener la misma coordenada vertical por lo que las cámaras poseen los mismos parámetros intrínsecos.

Se tiene entonces el plano de la imagen antiguo R_o y el nuevo R_n , que se obtendrá a través de la rectificación, por lo que a partir de las antiguas $\tilde{\mathbf{P}}_{oi}$ y $\tilde{\mathbf{P}}_{od}$ se necesita encontrar las nuevas $\tilde{\mathbf{P}}_{ni}$ y $\tilde{\mathbf{P}}_{nd}$, que serán, tal como se vio antes de la forma:

$$\tilde{\mathbf{P}}_{ni} = \mathbf{A}_{ni} (\mathbf{R}_{ni} | -\mathbf{R}_{ni} \mathbf{c}_i) \quad (\text{B.26})$$

La matriz de parámetros intrínsecos, $\mathbf{A}_{ni} = \mathbf{A}_{nd} = \mathbf{A}_n$, tiene que ser la misma para ambas cámaras ya que deben reflejar que son cámaras idénticas. Aunque teóricamente sus valores son arbitrarios, unos deformarán más la imagen que otros y pueden dar lugar a pequeños errores. En este caso se han tomado la media de los valores obtenidos en la calibración y se ha determinado que la distancia focal para el eje vertical y horizontal es la misma:

$$f_u = f_v = \frac{f_{ui} + f_{vi} + f_{ud} + f_{vd}}{4} u_o = \frac{u_{oi} + u_{od}}{2} v_o = \frac{v_{oi} + v_{od}}{2} \quad (\text{B.27})$$

Las matrices de rotación deben de ser idénticas para las dos cámaras, $\mathbf{R}_{ni} = \mathbf{R}_{nd} = \mathbf{R}_n$, ya que esto indica que ambos ejes ópticos están en la misma dirección. Esta última matriz se obtendrá fila a fila:

$$\mathbf{R}_n = \begin{pmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \mathbf{r}_3^T \end{pmatrix} \quad (\text{B.28})$$

Así el eje horizontal será aquel que contenga a los dos centros, que como deben de ser los mismos que en el sistema antiguo se obtienen a través de las ecuaciones (B.18) y (B.25):

$$\begin{aligned} c_i &= -\mathbf{P}_{oi}^{-1} \tilde{\mathbf{p}}_{oi} = \mathbf{0} \\ c_d &= -\mathbf{P}_{od}^{-1} \tilde{\mathbf{p}}_{od} = -\mathbf{R}^{-1} \mathbf{A}_{od}^{-1} \mathbf{A}_{od} \mathbf{t} = -\mathbf{R}^{-1} \mathbf{t} \end{aligned} \quad (\text{B.29})$$

Por lo que:

$$\mathbf{r}_1 = \frac{\mathbf{c}_d - \mathbf{c}_i}{\|\mathbf{c}_d - \mathbf{c}_i\|} = \frac{\mathbf{c}_d}{\|\mathbf{c}_d\|} \quad (\text{B.30})$$

El nuevo eje vertical debe de ser ortogonal al eje horizontal. De entre todos los que lo cumple, se puede tomar aquel que lo sea respecto al antiguo eje z de la imagen izquierda.

$$\mathbf{r}_2 = (kx\mathbf{r}_1) \quad (\text{B.31})$$

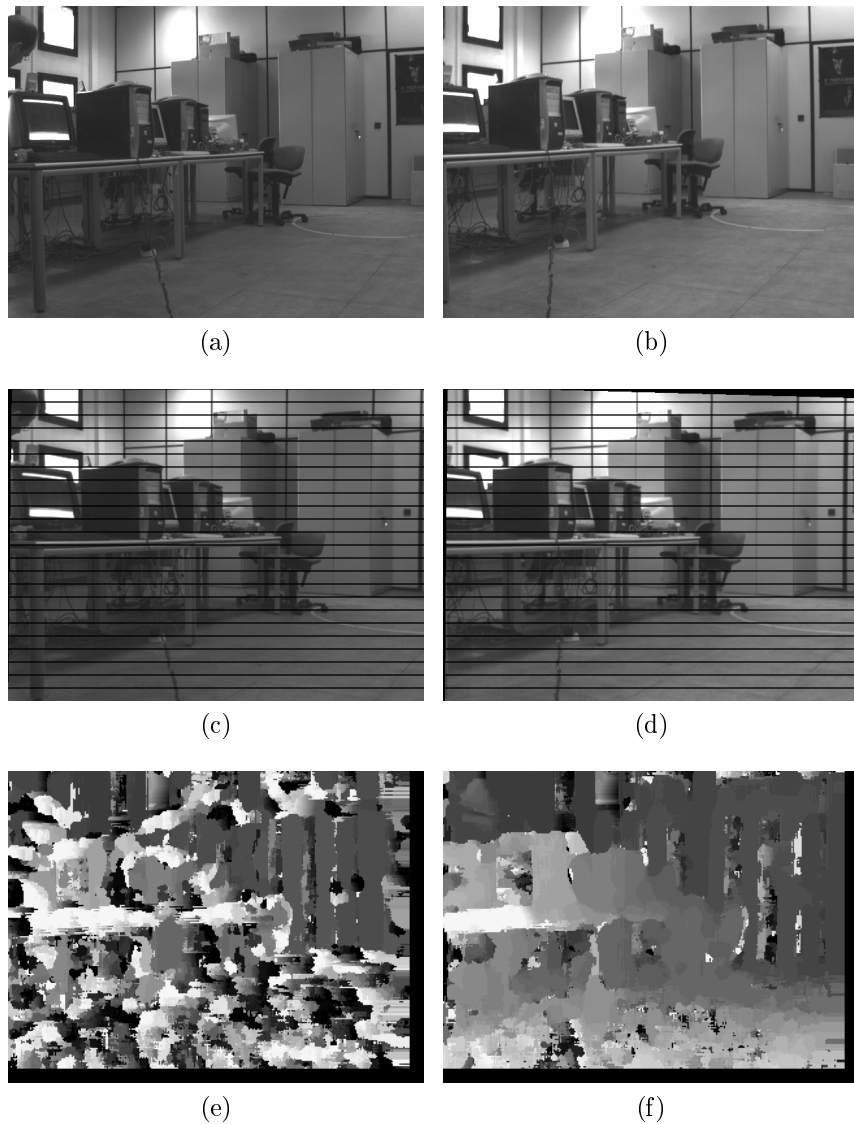


Figura B.11: Resultados; (a)(b) Imágenes izquierda y derecha sin rectificar; (c)(d) Imágenes izquierda y derecha rectificadas; (e) Mapa de disparidades entre (a) y (b); (f) Mapa de disparidades entre (c) y (d).

Por último, el nuevo eje z será:

$$\mathbf{r}_3 = (\mathbf{r}_1 \times \mathbf{r}_2) \quad (\text{B.32})$$

Se puede por tanto calcular:

$$\begin{aligned} \tilde{\mathbf{P}}_{ni} &= \mathbf{A}_n (\mathbf{R}_n | -\mathbf{R}_n \mathbf{c}_i) = \mathbf{A}_n (\mathbf{R}_n | \mathbf{0}) \\ \tilde{\mathbf{P}}_{nd} &= \mathbf{A}_n (\mathbf{R}_n | -\mathbf{R}_n \mathbf{c}_d) = \mathbf{A}_n (\mathbf{R}_n | \mathbf{R}_n \mathbf{R}^{-1} \mathbf{t}) \end{aligned} \quad (\text{B.33})$$

Solo queda obtener la transformación de rectificación, que para cada imagen es:

$$\mathbf{T} = \mathbf{P}_n \mathbf{P}_o^{-1} \quad (\text{B.34})$$

Esto se obtiene de la ecuación (B.20) donde:

$$\mathbf{M} = \mathbf{c}_i + \lambda_o \mathbf{P}_{oi}^{-1} \tilde{\mathbf{m}}_{oi} = \mathbf{c}_i + \lambda_n \mathbf{P}_{ni}^{-1} \tilde{\mathbf{m}}_{ni} \quad (\text{B.35})$$

Por lo que:

$$\tilde{\mathbf{m}}_{ni} = \lambda \mathbf{P}_{ni} \mathbf{P}_{oi}^{-1} \tilde{\mathbf{m}}_{oi} \quad (\text{B.36})$$

Y así:

$$\begin{aligned} \mathbf{T}_i &= \mathbf{P}_{ni} \mathbf{P}_{oi}^{-1} = \mathbf{A}_n \mathbf{R}_n \mathbf{A}_{oi}^{-1} \\ \mathbf{T}_d &= \mathbf{P}_{nd} \mathbf{P}_{od}^{-1} = \mathbf{A}_n \mathbf{R}_n \mathbf{R}^{-1} \mathbf{A}_{od}^{-1} \end{aligned} \quad (\text{B.37})$$

Para ver la importancia de la rectificación y los resultados, se puede observar la figura B.11. Las dos primeras imágenes no están rectificadas y se aprecia que no son totalmente paralelas. Aunque pueda pensarse que las diferencias son muy pequeñas, en figura B.11e se tiene el resultado de obtener el mapa de disparidades en el que se observa que los resultados erróneos. El contraste con los resultados mostrados en figura B.11f, a partir de las imágenes rectificadas es notable.

Apéndice C

Modelo parabólico de la carretera

C.1. Desarrollo en serie de Taylor de la circunferencia

Una circunferencia de centro (x_0, y_0) y radio R se expresa como:

$$(x - x_0)^2 + (y - y_0)^2 = R^2 \quad (\text{C.1})$$

Considerando la coordenada y como variable independiente (más cómodo de tratar a la hora de procesar las imágenes), el desarrollo en serie de Taylor de orden 2 en torno a $y = y_p$ será:

$$x(y) \approx x|_{y_p} + x'|_{y_p}(y - y_p) + \frac{x''}{2!} \Big|_{y_p} (y - y_p)^2 \quad (\text{C.2})$$

Despejando $x(y)$ en (C.1) y derivando se obtiene:

$$\begin{cases} x(y) = x_0 \pm \sqrt{R^2 - (y - y_0)^2} \\ x'(y) = \frac{\mp(y - y_0)}{\sqrt{R^2 - (y - y_0)^2}} \\ x''(y) = \frac{\mp R^2}{[R^2 - (y - y_0)^2]^{3/2}} \end{cases} \quad (\text{C.3})$$

Eligiendo una de las soluciones anteriores, e introduciendo la curvatura $C \equiv 1/R$, y suponiendo que $R^2 \gg (y_p - y_0)^2$, es decir, considerando que las curvaturas son pequeñas¹, las ecuaciones anteriores pueden reescribirse

¹Según la normativa 3.1-IC de trazado de carreteras (BOE, 2000), el radio mínimo para carreteras del grupo C-100 es de 250 m. En el peor caso, siendo y_p el punto más próximo al vehículo, e y_0 el más alejado, se tiene $y_p - y_0 = 40$ m. Por tanto, $(y_p - y_0)^2/R^2 = 40^2/250^2 = 0,0256 \ll 1$.

como:

$$\begin{cases} x(y)|_{y_p} \approx x_0 - R = x_0 + \frac{1}{C} \\ x'(y)|_{y_p} \approx \frac{y_p - y_0}{R} = (y_p - y_0) \cdot C \\ x''(y)|_{y_p} \approx \frac{1}{R} = C \end{cases} \quad (\text{C.4})$$

Sustituyendo (C.4) en (C.2), se tiene:

$$\begin{aligned} x &= x_0 - \frac{1}{C} + C(y_p - y_0)(y - y_0) + \frac{C}{2}(y - y_p)^2 = \\ &= x_0 - \frac{1}{C} + C(y_p - y_0)y - C(y_p - y_0)y_p + \frac{C}{2}y_p^2 + \\ &\quad - 2y_p \frac{C}{2}y + \frac{C}{2}y^2 = \\ &= \underbrace{\left(x_0 - \frac{1}{C} - C(y_p - y_0)y_p + \frac{C}{2}y_p^2 \right)}_b + \underbrace{(-Cy_0)}_m y + \frac{C}{2}y^2 = \\ &= b + m \cdot y + \frac{C}{2} \cdot y^2 \end{aligned} \quad (\text{C.5})$$

Particularizando (C.5) para $y_p = 0$, es decir, haciendo el desarrollo en torno al punto más próximo al vehículo, se obtiene una expresión considerablemente más sencilla:

$$x = \left(x_0 - \frac{1}{C} \right) - (C y_0) \cdot y + \frac{C}{2} \cdot y^2 \quad (\text{C.6})$$

C.2. Modelo de Carril parabólico

A partir de la ecuación anterior (C.6), en este apartado se deduce un conjunto de ecuaciones que puede representar todas las líneas viales de separación de carriles.

Dado un sistema de coordenadas fijo al vehículo y orientado respecto a él, y un centro de curvatura cuyas coordenadas cartesianas son (x_0, y_0) , y cuyas coordenadas polares son (R_c, ψ) (figura C.1), se cumple:

$$\begin{cases} x_0 = R_c \cos \psi \\ y_0 = R_c \sin \psi \end{cases} \quad (\text{C.7})$$

Si el vehículo no se separa excesivamente del trazado de la carretera, de modo que $\psi \approx 0$, y considerando $C = 1/R_c$, podemos aproximar las

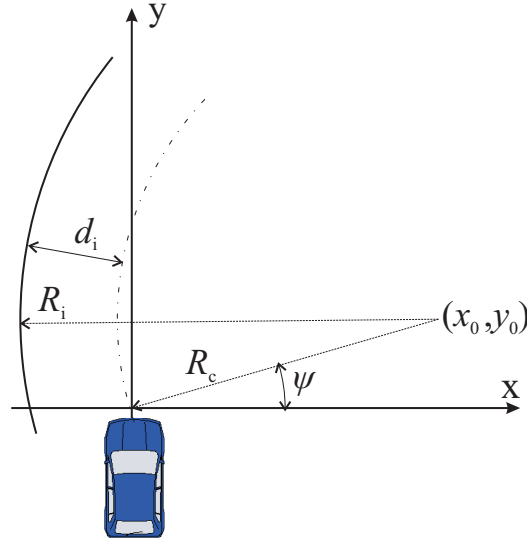


Figura C.1: Modelo parabólico de una línea vial

ecuaciones anteriores por:

$$\begin{cases} x_0 = R_c \cos \psi \approx \frac{1}{C} \\ y_0 = R_c \sin \psi \approx \frac{\psi}{C} \end{cases} \quad (\text{C.8})$$

Sea d_i la distancia del vehículo a una línea vial, y R_i su radio de curvatura (figura C.1), entonces $R_i = R_c + d_i$, que expresado en función de la curvatura $C_i = 1/R_i$, resulta:

$$C_i = \frac{C}{1 + d_i \cdot C} \quad (\text{C.9})$$

Sustituyendo (C.8) y (C.9) en (C.6), se obtiene:

$$x = -d_i - \frac{\psi}{1 + d_i \cdot C} y + \frac{C}{2(1 + d_i \cdot C)} y^2 \quad (\text{C.10})$$

Suponiendo que $d \cdot C \ll 1$, podemos aproximar la ecuación anterior por²:

$$x = -d_i - \psi \cdot y + \frac{C}{2} \cdot y^2 \quad (\text{C.11})$$

Tomando como referencia el eje central del carril, siendo d la distancia del vehículo dicho eje, W el ancho del carril, e i un índice que referencia las distintas líneas viales y que sólo toma valores impares (figura C.2), entonces

²Con el vehículo a la máxima distancia del borde del carril (unos 4 metros), y con la curvatura mínima en carreteras C-100: $d \cdot C = 4/250 = 0,016 \ll 1$.

$d_i = d + i \cdot W/2$. Sustituyendo esta expresión en (C.11), finalmente se obtiene:

$$x = \frac{C}{2} \cdot y^2 - \psi \cdot y - \left(d + i \cdot \frac{W}{2} \right); \quad i = \{\dots, -3, -1, 1, 3, \dots\} \quad (\text{C.12})$$

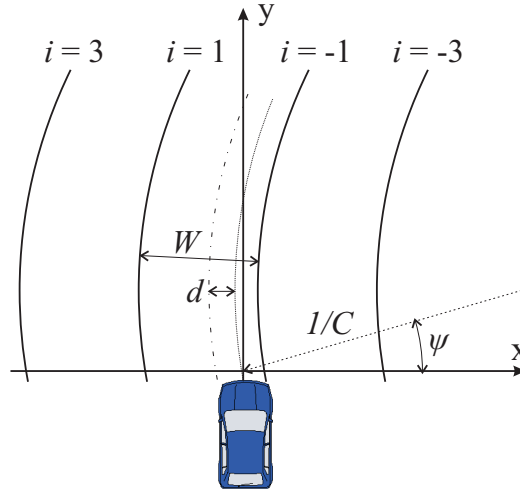


Figura C.2: Modelo parabólico de carril

Bibliografía

- APOSTOLOFF, N. y ZELINSKY, A. (2003). «Robust vision based lane tracking using multiple cues and particle filtering». En: *Computer Vision. IEEE International Symposium on*, pp. 558–563. doi: 10.1109/IVS.2003.1212973.
- ASSISTWARE TECHNOLOGY (2006). «SafeTRACK». [En línea] <http://www.assistware.com>. Última visita en: 3/4/2007.
- AUFRÉRE, R.; CHAPUIS, R. y CHAUSSE, F. (2001). «A model-driven approach for real-time road recognition». *Machine Vision and Applications*, **13(2)**, pp. 95–107.
- AYACHE, N. (1991). *Artificial Vision for Mobile Robots: Stereo Vision and Multisensory Perception*. The MIT Press.
- BAR-SHALOM, Y. y FORTMANN, T. E. (1988). *Tracking and Data Association*. Academic Press.
- BÜCHER, THOMAS (2000). «Measurement of Distance and Height in Images based on easy Attainable Calibration Parameters». En: *Proceedings IEEE Intelligent Vehicles Symposium*, pp. 314–319. Detroit, USA.
- BERTOZZI, M. y BROGGI, A. (1996). «Real-Time Lane and Obstacle Detection on the GOLD System». En: *Computer Vision. IEEE International Symposium on*, pp. 213–218. IEEE Computer Society, Tokyo, Japan.
- BERTOZZI, M.; BROGGI, A.; HILARIO, C.; FEDRIGA, R. I.; VEZZONI, G. y ROSE, M. DEL (2007). «Pedestrian Detection in Far Infrared Images based on the use of Probabilistic Templates». En: *IEEE Intelligent Vehicles Symposium*, Istanbul, Turkey.
- BOE (1987). «Norma 8.2-IC «Marcas viales», de la instrucción de carreteras». ORDEN de 16 de julio de 1987 (BOE núm. 185, de 4 ago. 1987, con corrección de errores en BOE núm. 233, de 29 sept. 1987).
- (2000). «Norma 3.1-IC. Trazado, de la Instrucción de Carreteras». ORDEN de 27 de diciembre de 1999 (BOE núm. 28, 2 feb. 2000), modificada parcialmente por la ORDEN de 13 septiembre 2001 (BOE núm. 231, 26 sept. 2001)..

- BOUGUET, JEAN-YVES (2006). «Camera Calibration Toolbox for Matlab». http://www.vision.caltech.edu/bouguetj/calib_doc
- BROGGI, A. (1995). «Robust Real-Time Lane and Road Detection in Critical Shadow Conditions». En: *Computer Vision. IEEE International Symposium on*, pp. 353–358. IEEE Computer Society, Coral Gables, Florida.
- BROGGI, A.; BERTOZZI, M. y FASCIOLI, A. (2001). «Self-calibration of a stereo vision system for automotive applications». En: *Proceedings 2001 IEEE International Conference on Robotics and Automation (ICRA)*, volumen 4, pp. 3698–3703. Seoul, South Korea.
- BROGGI, A.; BERTOZZI, M.; FASCIOLI, A. y CONTE, G. (1999). *Automatic Vehicle Guidance: The Experience of the ARGO Autonomous Vehicle*. World Scientific.
- BROWN, D.C. (1971). «Lens distortion for close-range Photogrammetry». *Photometric Engineering*, **37(8)**, pp. 855–866.
- CAMPBELL, N. W. y THOMAS, B. T. (1993). «Navigation of an Autonomous Road Vehicle using Lane Boundary Markings». En: D. Charnley (Ed.), *Intelligent Autonomous Vehicles. IFAC International Conference on*, pp. 169–174. Pergamon Press.
- CAMPBELL, N.W.; THOMAS, B.T. y TROSCIANKO, T. (1997). «Automatic segmentation and classification of outdoor images using neural networks». *International Journal of Neural Systems*, **8(1)**, pp. 137–44.
- CANNY, J. (1986). «A computational approach to edge detection». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8(6)**, pp. 679–698.
- CARRASCO, JUAN P.; ARMINGOL, JOSÉ M. y DE LA ESCALERA, ARTURO (2006). «Sistema de Ayuda a la Conducción para la supervisión de la Velocidad de un Vehículo». En: *XXVII Jornadas de Automática*, Almería.
- CARRASCO, JUAN P.; DE LA ESCALERA, ARTURO y ARMINGOL, JOSÉ M. (2008). «Driving Warning System Based on Visual Perception of Road Signs». En: *Third International Conference on Computer Vision Theory and Applications*, Madeira, Portugal.
- CARRASCO, JUAN PABLO; DE LA ESCALERA, ARTURO y ARMINGOL, JOSÉ M. (2007). «Speed Supervisor for Intelligent Vehicles». *Lecture Notes in Computer Science*, **4739**, pp. 1175–1182.
- CARRETEROS.ORG (1987). «Norma 8.2-IC». [En línea] http://www.carreteros.org/normativa/marcas_v/marcas_v.htm. Última visita en: 23/7/2007.

- CHAPUIS, R.; AUFRERE, R. y CHAUSSE, F. (2002). «Accurate road following and reconstruction by computer vision». *Intelligent Transportation Systems, IEEE Transactions on*, **3(4)**, pp. 261–270. doi: 10.1109/TITS.2002.804751.
- CHENG, HSU-YUNG; JENG, BOR-SHENN; TSENG, PEI-TING y FAN, KUO-CHIN (2006). «Lane Detection With Moving Vehicles in the Traffic Scenes». *Intelligent Transportation Systems. IEEE Transactions on*, **7(4)**, pp. 571–582. doi: 10.1109/TITS.2006.883940.
- CHIU, KUO-YU y LIN, SHENG-FUU (2005). «Lane Detecion using Color-Based Segmentation». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 705–710. Las Vegas, Nevada, U.S.A.. doi: 10.1109/IVS.2005.1505186.
- COLLADO, JUAN M.; HILARIO, CRISTINA; ARMINGOL, JOSÉ M^A y DE LA ESCALERA, ARTURO (2007). «On board camera perception and tracking of vehicles». En: *1st International Workshop on Robot Vision*, pp. 57–66. Barcelona, Spain.
- COLLADO, JUAN M.; HILARIO, CRISTINA; DE LA ESCALERA, ARTURO y ARMINGOL, JOSÉ M. (2004). «Model based vehicle detection for intelligent vehicles». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 572–577. Parma, Italy.
- (2005a). «Detección y clasificación de líneas viales mediante análisis frecuencial». En: *XXVI Jornadas de Automática*, Alicante.
- (2005b). «Detection and Classification of Road Lanes with a Frequency Analysis». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 78–83. Las Vegas, Nevada, USA.
- (2006). «Adaptative Road Lanes Detection and Classification». *Lecture Notes in Computer Science*, **4179**, pp. 1151–1162.
- COLLADO, JUAN M.; HILARIO, CRISTINA; DE LA ESCALERA, ARTURO y ARMINGOL, JOSE M. (2008). «Multi-Lane Visual Perception for Lane Departure Warning Systems». En: *3rd International Conference on Computer Vision Theory and Applications*, Funchal, Madeira (Portugal).
- CONTE, GIANNI; BROGGI, ALBERTO; BERTOZZI, MASSIMO; FASCIOLI, ALESSANDRA; PIAZZI, AURELIO y BIANCO, CORRADO GUARINO LO (1998). «MilleMiglia in Automatico». Última visita en: 3/4/2007.
<http://www.argo.ce.unipr.it>
- CONTINENTAL AUTOMOTIVE (2009). «Lane Departure Warning / Lane Keeping System (LDW / LKS)». [En línea] http://www.conti-online.com/generator/www/de/en/continental/automotive/themes/commercial_

- [vehicles/safety/adas/ldw/ldw_lks_en.html](#). Última visita en 16/07/2009.
- CORKE, P.I.; SYMEONIDIS, D. y USHER, K. (2003). «Tracking road edges in the panospheric image plane». En: *Intelligent Robots and Systems (IROS). IEEE/RSJ International Conference on*, volumen 2, pp. 1330–1335.
- CRISMAN, J. D. y THORPE, C. E. (1991). «UNSCARF-a color vision system for the detection of unstructured roads». En: *Robotics and Automation. IEEE International Conference on*, volumen 3, pp. 2496–2501. Sacramento, CA, USA. doi: 10.1109/ROBOT.1991.132000.
- (1993). «SCARF: A color vision system that tracks roads and intersections». *Robotics and Automation. IEEE Transactions on*, **9**, pp. 49–58. doi: 10.1109/70.210794.
- DANG, T. y HOFFMANN, C. (2004). «Stereo calibration in vehicles». En: *IEEE Intelligent Vehicles Symposium*, pp. 268–273. Parma, Italy.
- DANKERS, ANDREW y ZELINSKY, ALEXANDER (2004). «CeDAR: A real-world vision system. Mechanism, control and visual processing». *Machine Vision and Applications*, **16(1)**, pp. 47–58. doi: 10.1007/s00138-004-0156-3.
- DE LA ESCALERA, A. (1995). *Percepción del entorno y localización de robots móviles autónomos*. Tesis doctoral, Universidad Carlos III de Madrid, Leganés, Spain. Directores: E. Andrés Puente, L. Moreno Lorente.
- DE LA ESCALERA, A.; ARMINGOL, J. M^a y SALICHS, M. A. (2001). «Traffic Sign Detection for Driver Support Systems». En: *International Conference on Field and Service Robotics*, Espoo, Finland.
- DE LA ESCALERA, A.; ARMINGOL, JOSÉ M.; RODRIGUEZ, FRANCISCO J. y PASTOR, JOSE M. (2004). «Visual sign information extraction and identification by deformable models for intelligent vehicles». *Intelligent Transportation Systems, IEEE Transactions on*, **5(2)**, pp. 57–68.
- DE LA ESCALERA, ARTURO; ARMINGOL, JOSÉ M. y MATA, MARIO (2003). «Traffic Sign Recognition and Analysis for Intelligent Vehicles». *Image and Vision Computing*, **11(3)**, pp. 247–258.
- DE LA ESCALERA, ARTURO; MORENO, LUIS; SALICHS, MIGUEL A. y ARMINGOL, JOSÉ M. (1994a). «Traffic Sign Perception in Road Environments». En: *International Symposium on Automotive Technology and Automation*, Aachen, Alemania.
- DE LA ESCALERA, ARTURO; MORENO, LUIS; SALICHS, MIGUEL A. y ARMINGOL., JOSÉ M. (1997). «Road Traffic Sign Detection and Classification». *IEEE Transactions on Industrial Electronics*, **44(6)**, pp. 848–859.

- DE LA ESCALERA, ARTURO; SALICHS, MIGUEL A.; MORENO, LUIS y PUENTE, E. A. (1994b). «Neural Traffic Signal Recognition for Autonomous Vehicles». En: *IEEE International Conference on Industrial Electronics, Control, Instrumentation and Automation*, Bologna, Italy.
- DEMENTHON, D. (1987). «A Zero-bank Algorithm for Inverse Perspective of a Road from a Single Image». En: *Robotics and Automation, IEEE International Conference on*, pp. 1444–1449. Raleigh, NC.
- DGT (2006a). «Anuario de accidentes, 2006». *Informe técnico*, Dirección General de Tráfico.
<http://www.dgt.es>
- (2006b). «Opiniones sobre Seguridad Vial de los Ciudadanos Españoles». *Informe técnico*, Observatorio Nacional de Seguridad Vial, Dirección General de Tráfico.
<http://www.dgt.es>
- (2006c). «Principales cifras de siniestralidad vial». *Informe técnico*, Dirección General de Tráfico, Observatorio Nacional de Seguridad Vial.
<http://www.dgt.es>
- DICKMANN, E. D. y MYSLIWETZ, B. D. (1992). «Recursive 3-D Road and Relative Ego-State Recognition». *Pattern Analysis and Machine Intelligence. IEEE Transactions on*, **14**(2), pp. 199–213. doi: 10.1109/34.121789.
- DOUCET, ARNAUD; DE FREITAS, NANDO y GORDON, NEIL (2001). *Sequential Monte Carlo methods in practice*. Springer.
- ERNST, S.; STILLER, C.; GOLDBECK, J. y ROESSIG, C. (1999). «Camera calibration for lane and obstacle detection». En: *Proceedings 1999 IEEE/IEEEJ/JSAP International Conference on Intelligent Transportation Systems*, pp. 356–361. Tokyo, Japan.
- FAUGERAS, O. (1993). *Three-dimensional computer vision: a geometric viewpoint*. MIT Press.
- FLORES, MARCO J.; ARMINGOL, JOSÉ M. y DE LA ESCALERA, ARTURO (2007). «New Probability Models for Face Detection and Tracking in Color Images». En: *International Symposium on Intelligent Signal Processing*, Alcalá de Henares, Spain.
- FODA, S.G. y DAWOUD, A.K. (2001). «Highway lane boundary determination for autonomous navigation». En: *Communications, Computers and signal Processing, 2001. PACRIM. 2001 IEEE Pacific Rim Conference on*, volumen 2, pp. 698–702. doi: 10.1109/PACRIM.2001.953728.

- FORESTI, G. L. y REGAZZONI, C. (1999). «3D road scene interpretation for autonomous vehicle driving». *Journal of Computing and Information Technology (CIT)*, **7**(4), pp. 277–294.
- FRANKE, U.; LOOSE, H. y KNOPPEL, C. (2007). «Lane Recognition on Country Roads». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 99–104. ISSN 1931-0587. doi: 10.1109/IVS.2007.4290098.
- FUSIELLO, A.; TRUCCO, E. y VERRI, A. (2000). «A compact algorithm for rectification of stereo pairs». *Machine Vision and Applications*, **12**, pp. 16–22.
- GLASER, SEBASTIEN; LABAYRADE, RAPHAËL; MAMMAR, SAÏD; DOURET, JEROME y LUSETTI, BENOIT (2006). «Validation of a Vision Based Time to Line Crossing Computation». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 200–205. Tokyo, Japan.
- GOLDBERG, DAVID E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Pub. Co..
- GONZALEZ, J.P. y OZGUNER, U. (2000). «Lane detection using histogram-based segmentation and decision trees». En: *Intelligent Transportation Systems. IEEE International Conference on*, pp. 346–351. Piscataway, NJ, USA. doi: 10.1109/ITSC.2000.881084.
- GREGOR, R.; LUTZELER, M.; PELLKOEFER, M.; SIEDERSBERGER, K.-H. y DICKMANN, E.D. (2002). «EMS-Vision: a perceptual system for autonomous vehicles». *Intelligent Transportation Systems, IEEE Transactions on*, **3**(1), pp. 48–59. doi: 10.1109/6979.994795.
- GUIDUCCI, A. (1999). «Parametric Model of the Perspective Projection of a Road with Applications to Lane Keeping and 3D Road Reconstruction». *Computer Vision and Image Understanding*, **73**(3), pp. 414–427.
- HARTLEY, R. y GUPTA, R. (1993). «Computing matched-epipolar projections». En: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 549–555.
- HELLA (2007). «Lane Departure Warning». [En línea] http://www.hella.com/produktion/HellaCOM/WebSite/Channels/AutoIndustry/Electronics/DA_Systems/DA_Systems.jsp. Última visita en: 10/10/2007.
- HILARIO, C.; COLLADO, J.M.; ARMINGOL, J.M. y DE LA ESCALERA, A. (2005a). «Pyramidal Image Analysis for Vehicle Detection». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 87–92. Las Vegas, Nevada, U.S.A..

- HILARIO, CRISTINA; COLLADO, JUAN M.; ARMINGOL, JOSÉ M. y DE LA ESCALERA, ARTURO (2005b). «Detección de peatones para vehículos inteligentes basada en modelos de contornos activos y visión estéreo». En: *XXVI Jornadas de Automática*, Alicante.
- (2005c). «Multi-Resolution Image Analysis for Vehicle Detection». En: *Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA). Proceedings of the 2nd*, volumen 1 de *Lecture Notes in Computer Science*, Vol. 3522, pp. 87–92. Springer, Estoril, Portugal.
- (2005d). «Multi-Resolution Image Analysis for Vehicle Detection». *Lecture Notes in Computer Science*, **3522**, pp. 579–586.
- (2005e). «Pedestrian Detection for Intelligent Vehicles based on Active Contour Models and Stereo Vision». En: *Computer Aided Systems Theory. 10th International Workshop on*, pp. 265–268. Gran Canaria, Spain.
- (2005f). «Pedestrian Detection for Intelligent Vehicles based on Active Contour Models and Stereo Vision». *Lecture Notes in Computer Science*, **3643**, pp. 537–542.
- (2006). «Visual Perception and Tracking of Vehicles for Driver Assistance Systems». En: *Intelligent Vehicles Symposium, 2006 IEEE*, pp. 94–99. Tokyo, Japan.
- HILARIO, CRISTINA; COLLADO, JUAN M.; DE LA ESCALERA, ARTURO y ARMINGOL, JOSÉ M. (2004). «Driver Assistance System Based on Computer Vision for Vehicle Detection». En: *Intelligent Autonomous Vehicles. Proceedings of the 5th IFAC/EURON Symposium on*, Lisboa, Portugal.
- HONDA (2009). «Sistema LKAS del Honda Accord». [En línea] <http://www.honda.es>. Se puede acceder siguiendo la ruta: Gama -> Accord -> Tecnología Avanzada -> Seguridad Activa -> sistema LKAS. Última visita en: 17/7/2009.
- HOUGH, PAUL V.C. (1959). «Machine Analysis of Bubble Chamber Pictures». En: CERN (Ed.), *High Energy Accelerators and Instrumentation. International Conference on*, pp. 554–556.
- IENG, SIO-SONG; TAREL, JEAN-PHILIPPE y LABAYRADE, R. (2003). «On the design of a single lane-markings detectors regardless the on-board camera's position». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 564–569. Columbus, OH, USA. doi: 10.1109/IVS.2003.1212974.
- IRISBUSIVECO (2009). «Civis». [En línea] <http://irisbusiveco.com/prod/civis.html>. Última visita en: 17/7/2009.

- ISARD, M. y BLAKE, A. (1998). «CONDENSATION – Conditional density propagation for visual tracking». *International Journal of Computer Vision*, **29**(1), pp. 5–28. Kluwer Academic Publishers.
<http://search.epnet.com/direct.asp?an=6065799&db=inh>
- ISARD, MICHAEL ACHESON (1998). *Visual Motion Analysis by Probabilistic Propagation of Conditional Density*. Tesis doctoral, Oxford University.
- ITERIS INC. (2000). «AutoVue, Lane Departure Warning System». [En línea] <http://www.iteris.com/av/avindex.html>. Última visita en: 11-10-2007.
- JEONG, P. y NEDEVSKI, S. (2003). «Intelligent road detection based on local averaging classifier in real-time environments». En: *Image Analysis and Processing. Proceedings of the 12th International Conference on*, pp. 245–249. doi: 10.1109/ICIAP.2003.1234057.
- JEONG, PANGYU y NEDEVSKI, S. (2005). «Efficient and robust classification method using combined feature vector for lane detection». *Circuits and Systems for Video Technology, IEEE Transactions on*, **15**(4), pp. 528–537. doi: 10.1109/TCSVT.2005.844453.
- JOCHER, T. y POMERLEAU, D. (1995). «No Hands Across America Navlab 5 USA tour». http://www.cs.cmu.edu/afs/cs/user/tjocher/www/nhaa/nhaa_home_page.html. Última visita en: 2/4/2007.
- JUNG, CLAUDIO ROSITO y KELBER, CHRISTIAN ROBERTO (2005). «Lane following and lane departure using a linear-parabolic model». *Image and Vision Computing*, **23**(13), pp. 1192–1202. doi: 10.1016/j.imavis.2005.07.018.
- KANATANI, K. y WATANABE, K. (1990). «Reconstruction of 3-D Road Geometry from Images for Autonomous Land Vehicles». *Robotics and Automation. IEEE Transactions on*, **6**(1). doi: 10.1109/70.88128.
- KIM, ZU (2006). «Realtime lane tracking of curved local road». En: *Intelligent Transportation Systems, IEEE International Conference on*, pp. 1149–1155. doi: 10.1109/ITSC.2006.1707377.
- KLUGE, K. y LAKSHMANAN, S. (1995). «A deformable-template approach to lane detection». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 54–59. doi: 10.1109/IVS.1995.528257.
- KLUGE, KARL (1993). *YARF: An Open-Ended Framework for Robot Road Following*. Tesis doctoral, Carnegie-Mellon University, School of Computer Science.
<http://handle.dtic.mil/100.2/ADA266788>

- KREUCHER, C. y LAKSHMANAN, S. (1999). «LANA: A Lane Extraction Algorithm that uses Frequency Domain Features». *Robotics and Automation. IEEE Transactions on*, **15**(2), pp. 343–350.
- LABAYRADE, R. y AUBERT, D. (2003). «A single framework for vehicle roll, pitch, yaw estimation and obstacles detection by stereovision». En: *Proceedings IEEE Intelligent Vehicles Symposium*, pp. 31–36.
- LABAYRADE, R.; DOURET, J. y AUBERT, D. (2006). «A multi-model lane detector that handles road singularities». En: *Intelligent Transportation Systems, IEEE International Conference on*, pp. 1143–1148. doi: 10.1109/ITSC.2006.1707376.
- LEAVERS, V.F. (1992). *Shape detection in computer vision using the hough transform*.
- LEE, H.-J. y DENG, C.-T. (1991). «Determining camera models from multiple frames». *Journal of Information Science and Engineering*, **12**(2), pp. 193–214.
- LEI, GUO; JIANQIANG, WANG y KEQIANG, LI (2006). «Lane Keeping System Based on THASV-II Platform». En: *Vehicular Electronics and Safety, 2006. ICVES 2006. IEEE International Conference on*, pp. 305–308. doi: 10.1109/ICVES.2006.371604.
- LI, QING; ZHENG, NANNING y CHENG, HONG (2004). «Springrobot: a prototype autonomous vehicle and its algorithms for lane detection». *Intelligent Transportation Systems, IEEE Transactions on*, **5**(4), pp. 300–308. doi: 10.1109/TITS.2004.838220.
- LIN, X. y CHEN, S. (1991). «Color image Segmentation Using Modified HSI System for Road Following». En: *International Conference on Robotics and Automation. Proceedings of the IEEE*, Sacramento, California, EEUU. doi: 10.1109/ROBOT.1991.131921.
- LUETZELER, M. y BATEN, S. (2000). «Road recognition for a tracked vehicle». En: *Conference on Enhanced and Synthetic Vision. Proceedings of the SPIE*, volumen 4023, pp. 171–180. Orlando (FL), USA. doi: 10.1117/12.389340.
- MAURER, M.; MAURER, M.; BEHRINGER, R.; FURST, S.; THOMANEK, F. y DICKMANN, E.D. (1996). «A compact vision system for road vehicle guidance». En: R. Behringer (Ed.), *International Conference on Pattern Recognition*, volumen 3, pp. 313–317. doi: 10.1109/ICPR.1996.546962.
- MCCALL, J.C. y TRIVEDI, M.M. (2006). «Video-based lane estimation and tracking for driver assistance: survey, system, and evaluation». *Intelligent Transportation Systems, IEEE Transactions on*, **7**(1), pp. 20–37. doi: 10.1109/TITS.2006.869595.

MOBILEYE. <http://www.mobileye-vision.com/>.

MORGENTHALER, D.G.; HENNESSY, S. y DEMENTHON, D. (1990). «Range-video fusion and comparison of inverse perspective algorithms in static images». *Systems, Man and Cybernetics. IEEE Transactions on*, **20**(6), pp. 1301–1312. doi: 10.1109/21.61202.

NEDEVSCI, S.; DANESCU, R.; MARITA, T.; ONIGA, F.; POCOL, C.; SOBEL, S.; GRAF, T. y SCHMIDT, R. (2005). «Driving environment perception using stereovision». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 331–336. Las Vegas, Nevada, U.S.A.. doi: 10.1109/IVS.2005.1505124.

NEDEVSCI, S.; SCHMIDT, R.; GRAF, T.; DANESCU, R.; FRENIU, D.; MARITA, T.; ONIGA, F. y POCOL, C. (2004). «3D lane detection system based on stereovision». En: *Intelligent Transportation Systems, IEEE International Conference on*, pp. 161–166. Washington, D.C., USA.

NEDEVSCI, S.; VANCEA, C.; MARITA, T. y GRAF, T. (2006). «On-line calibration method for stereovision systems used in vehicle applications». En: *Intelligent Transportation Systems Conference. Proceedings of the*, pp. 957–962. doi: 10.1109/ITSC.2006.1706868.

PALMER, P. L.; KITTLER, J. y PETROU, M. (1993a). «An Optimizing Line Finder Using a Hough Transform Algorithm». *Computer Vision and Image Understanding*, **68**(1), pp. 1–23.

PALMER, P. L.; PETROU, M. y KITTLER, J. (1993b). «A Hough Transform Algorithm with a 2D Hypothesis Testing Kernel». *CVGIP: Image Understanding*, **58**, pp. 221–234.

PAPADIMITRIOU, D. V. y DENNIS, T. J. (1996). «Epipolar line estimation and rectification for stereo images pairs». *IEEE Transactions on Image Processing*, **5**(4), pp. 672–676.

PARK, JONG WOONG; LEE, JOON WOONG y JHANG, KYUNG YOUNG (2003). «A lane-curve detection based on an LCF». *Pattern Recognition Letters*, **24**(14), pp. 2301–2313. ISSN 0167-8655. doi: [http://dx.doi.org/10.1016/S0167-8655\(03\)00056-4](http://dx.doi.org/10.1016/S0167-8655(03)00056-4).

POMERLEAU, D. y JOCHEM, T. (1996). «Rapidly adapting machine vision for automated vehicle steering». Traducido por english, *IEEE Expert*, **11**(2), pp. 19–27. doi: 10.1109/64.491277.

POMERLEAU, D.A. (1993). *Neural Network Perception for Mobile Robot Guidance*. Kluwer Academic Publishers.

- RICHTER, S. y WETZEL, D. (1995). «Adaptive change of strategies for road recognition and road tracking». En: *Intelligence Autonomous Systems. IAS-4. Proceedings of the International Conference*, Karlsruhe, Germany.
- RISACK, R.; KLAUSMANN, P.; KÜGER, W. y W. ENKELMANN (1998). «Robust lane recognition embedded in a real-time driver assistance system». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 35–40.
- ROBERT, L.; ZELLER, C.; FAUGERAS, O. y HÉBERT, M. (1997). *Visual Navigation: From Biological Systems to Unmanned Ground Vehicles*. capítulo 5: Applications of non-metric vision to some visually-guided robotics tasks, pp. 89–134.
- SAPPA, ANGEL D.; GERÓNIMO, DAVID; DORNAIKA, FADI y LÓPEZ, ANTONIO (2006). «Real time vehicle pose using on-board stereo vision system». En: *Image Analysis and Recognition (ICIAR 2006). Third International Conference on*, pp. 205–216. Pova De Varzim, Portugal. doi: 10.1007/11867661.
- SCHUMACHER, ERNST FRIEDRICH (1973). *Lo pequeño es hermoso*.
- SHANAHAN, J.; THOMAS, B.; M. MIRMEHDI; CAMPBELL, N.; MARTÍN, T. y BALDWIN, J. (1999). «Road recognition using fuzzy classifiers». En: *10th British Machine Vision Conference, Proceedings of the*, volumen 2, pp. 432–442. Nottingham, UK.
- SOTELO, MIGUEL A.; RODRIGUEZ, FRANCISCO J.; MAGDALENA, LUIS; BERGASA, LUIS M. y BOQUETE, LUCIANO (2004). «A Color Vision-Based Lane Tracking System for Autonomous Driving on Unmarked Roads». *Autonomous Robots*, **16**(1), pp. 95–116. doi: 10.1023/B:AURO.00000008673.96984.28.
- SOUTHALL, B. y TAYLOR, C.J. (2001). «Stochastic road shape estimation». En: *Computer Vision (ICCV). Proceedings of the 8th IEEE International Conference on*, volumen 1, pp. 205–212.
- SUN, TSUNG-YING; TSAI, SHANG-JENG y CHAN, V. (2006). «HSI color model based lane-marking detection». En: *Intelligent Transportation Systems. IEEE International Conference on*, pp. 1168–1172. doi: 10.1109/ITSC.2006.1707380.
- SUTTORP, T. y BUCHER, T. (2006). «Robust vanishing point estimation for driver assistance». En: *Intelligent Transportation Systems. IEEE International Conference on*, pp. 1550–1555. doi: 10.1109/ITSC.2006.1707444.
- TAOKA, T.; MANABE, M. y FUKUI, M. (2007). «An Efficient Curvature Lane Recognition Algorithm by Piecewise Linear Approach». En: *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, pp. 2530–2534. ISSN 1550-2252.

- TAYLOR, C.J.; MALIK, J. y WEBER, J. (1996). «A real-time approach to stereopsis and lane-finding». En: *Intelligent Vehicles Symposium, Proceedings of the IEEE*, pp. 207–212. doi: 10.1109/IVS.1996.566379.
- THORPE, C.; HEBERT, M.; KANADE, T. y SHAFER, S. (1991). «Toward Autonomous Driving: The CMU Navlab». *IEEE Expert*, **6**(4), pp. 31–42. doi: 10.1109/64.85919.
- TOYOTA (2007). «Lexus Range LS 460 - Lane Keeping Assist». [En línea] <http://www.lexus-europe.com>. [En línea]: Última visita en: 10/10/2007.
- TRIVEDI, M. M.; GANDHI, T. y MCCALL, J. (2007). «Looking-In and Looking-Out of a Vehicle: Computer-Vision-Based Enhanced Vehicle Safety». *Intelligent Transportation Systems, IEEE Transactions on*, **8**(1), pp. 108–120. doi: 10.1109/TITS.2006.889442.
- TRW AUTOMOTIVE. [En línea] <http://www.trwauto.com>. Última visita en: 3/4/2007. Ruta de acceso: TRW Home > Technology Portfolio > Active Safety Systems > Driver Assist Systems > Lane Guide Systems..
- TURCHETTO, R. y MANDUCHI, R. (2003). «Visual curb localization for autonomous navigation». En: *Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, volumen 2, pp. 1336–1342.
- TURK, M.A.; MORGENTHALER, D.G.; GREMBAN, K.D. y MARRA, M. (1988). «VITS - A vision system for autonomous land vehicle navigation». *Pattern Analysis and Machine Intelligence. IEEE Transactions on*, **10**, pp. 342–361. doi: 10.1109/34.3899.
- VLACIC, L.; PARENT, M. y HARASHIMA, F. (2001). *Intelligent Vehicle Technologies: Theory And Applications*. Butterworth-Heinemann.
- VOISIN, VINCENT; AVILA, MANUEL; EMILE, BRUNO; BÉGOT, STEPHANE y BARDET, JEAN-CHRISTOPHE (2005). «Road Markings Detection and Tracking Using Hough Transform and Kalman Filter.» En: *Advanced Concepts for Intelligent Vision Systems (ACIVS 2005), Proceedings of the 7th International Conference on*, volumen 3708 de *Lecture Notes in Computer Science*, pp. 76–83. Springer, Antwerp, Belgium.
- WANG, HONG y CHEN, QIANG (2006). «Real-time lane detection in various conditions and night cases». En: *Intelligent Transportation Systems. IEEE International Conference on*, pp. 1226–1231. doi: 10.1109/ITSC.2006.1707390.
- WANG, Y.; TEOH, E. K. y SHEN, D. (2004). «Lane detection and tracking using B-Snake». *Image and Vision Computing*, **22**, pp. 269–280.

- WANG, YUE; SHEN, DINGGANG y TEOH, EAM KHWANG (2000). «Lane detection using spline model». *Pattern Recognition Letters*, **21**(8), pp. 677–689. *Pattern Recognition Letters*, vol.21, no.8, July 2000. p. 677-689..
- WATANABE, A. y NISHIDA, M. (2005). «Lane Detection for a Steering Assistance System». En: *Intelligent Vehicles Symposium. Proceedings of the IEEE*, pp. 159–164. Las Vegas, Nevada, U.S.A.. doi: 10.1109/IVS.2005.1505095.
- WILLIAMSON, T.; WILLIAMSON, T. y THORPE, C. (1999). «A trinocular stereo system for highway obstacle detection». En: C. Thorpe (Ed.), *Proc. IEEE International Conference on Robotics and Automation*, volumen 3, pp. 2267–2273. doi: 10.1109/ROBOT.1999.770443.
- YUILLE, A. L. y COUGHLAN, J. M. (2000). «Fundamental limits of Bayesian inference: order parameters and phase transitions for road tracking». *Intelligent Transportation Systems, IEEE Transactions on*, **22**(2), pp. 160–173.
- ZHANG, Z. (1999). «Flexible Camera Calibration By Viewing a Plane From Unknown Orientations». En: *International Conference on Computer Vision (ICCV'99)*, pp. 666–673. Corfu, Greece.
- (2000). «A Flexible New Technique for Camera Calibration». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(11), pp. 1330–1334.
- ZHOU, YONG; XU, RONG; HU, XIAOFENG y YE, QINGTAI (2006). «A robust lane detection and tracking method based on computer vision». *Measurement Science and Technology*, **17**(4), pp. 736–745. doi: 10.1088/0957-0233/17/4/020.
- ZHU, ZHIGANG; XI, HAOJUN y XU, GUANGYOU (1996). «Combining rotation-invariance images and neural networks for road scene understanding». En: *Neural Networks. IEEE International Conference on*, volumen 3, pp. 1732–1737. doi: 10.1109/ICNN.1996.549162.